

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2005, 22:00:09 ; Search time 9493.76 Seconds
(without alignment)
11376.592 Million cell updates/sec

Title: US-09-622-964A-2
Perfect score: 2229
Sequence: 1 caggagatccaccagccta.....aaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 2422767955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2229	100.0	2229	6	BD136720 Best's ma
2	2208.4	99.1	2210	9	AF057169 Homo sapi
3	2166.2	97.2	2171	9	AF073501 Homo sapi
4	2155.8	96.7	2170	6	CQ729633 Sequence
5	2022.6	90.7	2187	9	AY357925 Macaca fa
6	2013	90.3	2429	6	BD136721 Best's ma
7	2002.4	89.8	2420	9	AF057170 Homo sapi
8	1793	80.4	2441	9	BC015220 Homo sapi
9	1756.4	78.8	1758	9	AY515704 Homo sapi
10	1754.8	78.7	1758	6	AX745964 Sequence
11	1584	71.1	2435	9	BC041664 Homo sapi
12	1247.2	56.0	1263	6	BD132659 Secretd
13	925.8	41.5	2035	10	BC079048 Rattus no
14	893.4	40.1	1904	10	AY450427 Mus muscu
15	871.2	39.1	1916	6	BD136744 Best's ma
16	861.4	38.6	1289	6	AY064707 Sus scrof
17	642.2	28.8	16125	6	BD136719 Best's ma
18	642.2	28.8	142092	9	AF139813 Homo sapi
19	642.2	28.8	163024	9	AF006260 Homo sapi

c	20	642.2	28.8	163915	2	AC087451	AC087451 Homo sapi
	21	642.2	28.8	166867	9	AP003733	AP003733 Homo sapi
	22	642.2	28.8	196080	9	AC004228	AC004228 Homo sapi
	23	640.6	28.7	706	9	HSVM2P10	AF073499 Homo sapi
	24	640.6	28.7	112309	9	AC003025	AC003025 Human chr
	25	640.6	28.7	133683	2	AC084857	AC084857 Homo sapi
	26	599	26.9	726	9	AY357926	AY357926 Macaca fa
	27	577.2	25.9	2137	6	AX745966	AX745966 Sequence
	28	576.8	25.9	1530	6	CQ731444	CQ731444 Sequence
	29	576.8	25.9	1530	6	AX745993	AX745993 Sequence
	30	576.8	25.9	1530	9	AY515705	AY515705 Homo sapi
	31	576.8	25.9	1908	9	AF440756	AF440756 Homo sapi
	32	552.8	24.8	160169	2	AC051664	AC051664 Homo sapi
c	33	538.4	24.2	1956	10	BC031186	BC031186 Mus muscu
	34	538.4	24.2	1957	10	BC019528	BC019528 Mus muscu
	35	538.4	24.2	1957	10	BC036157	BC036157 Mus muscu
	36	538.4	24.2	2028	10	BC036163	BC036163 Mus muscu
	37	535.2	24.0	1954	10	AY450428	AY450428 Mus muscu
	38	535	24.0	1861	9	AK000139	AK000139 Homo sapi
	39	494.4	22.2	1898	5	BC043854	BC043854 Xenopus l
	40	494.4	22.2	1921	5	AY273825	AY273825 Xenopus l
	41	484.6	21.7	1882	5	CR760914	CR760914 Xenopus t
	42	484.6	21.7	2021	5	BC061379	BC061379 Xenopus t
	43	478.4	21.5	2064	5	BC084229	BC084229 Xenopus t
	44	474	21.3	2719	10	AY450426	AY450426 Mus muscu
	45	472	21.2	1650	5	AY273826	AY273826 Xenopus l

ALIGNMENTS

RESULT 1
BD136720
LOCUS BD136720
DEFINITION Best's macular dystrophy gene.
ACCESSION BD136720.1 GI:23231665
VERSION JP 2002504559-A/2.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens

REFERENCE
AUTHORS Petrukhin,K., Caskey,T.C., Metzker,M. and Wadelius,C.
TITLE Best's macular dystrophy gene
JOURNAL Patent: JP 2002504559-A 2 12-FEB-2002;
COMMENT MERCK & CO INC, CLAES WADELIVS
OS Homo sapiens (human)
PN JP 2002504559-A/2
PP 22-FEB-1999 JP 2000533447
PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUKHIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI
WADELIVS

PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC
C12P21/08,
PC C12N5/00, C12N15/00
CC Best's macular dystrophy gene
FH Key location/Qualifiers
FT source 1..2229
FT /organism="Homo sapiens (human)".

FEATURES

source

location/Qualifiers
1..2229
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2229; DB 6; Length 2229;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
oy 1 CAGGAGTCCACAGCCTAGTCCAGACCTCTGTGGATCATCGACCCACCTGGAA 60

Db 1 CAGGAGTCCCAAGCTAGTCCGAGACCTTCTGAGGATATGGAGCCACCTGGAA 60
Qy 61 CCCACCTGACCCCAAGCCCACTGCTGACGCCCACTGCTGGCATGACCATCATTA 120
Db 61 CCCACCTGACCCCAAGCCCACTGCTGACGCCCACTGCTGGCATGACCATCATTA 120
Qy 121 CAGCCAAAGTGGTAAATGCCCCGTTAGGCTCTTCCGCGCTGCTGCTGGCGGG 180
Db 121 CAGCCAAAGTGGTAAATGCCCCGTTAGGCTCTTCCGCGCTGCTGCTGGCGGG 180
Qy 181 GCGAGCTTCAAGCTGCTATATGCGAGATTCTTAATCTTCTGCTGCTCACTCA 240
Db 181 GCGAGCTTCAAGCTGCTATATGCGAGATTCTTAATCTTCTGCTGCTCACTCA 240
Qy 241 TCGGCTTTATTTATAGGCTGGCCCTCAGAGAAACAAGCTGATGTTTGAAGAACTGA 300
Db 241 TCGGCTTTATTTATAGGCTGGCCCTCAGAGAAACAAGCTGATGTTTGAAGAACTGA 300
Qy 301 CTCTGATTTGCGACAGCTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCA 360
Db 301 CTCTGATTTGCGACAGCTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCA 360
Qy 361 TGAAGCTGGTGGTGAACCCGCTGGTGAACCAATGAGAACTTGGCCGCGACCGC 420
Db 361 TGAAGCTGGTGGTGAACCCGCTGGTGAACCAATGAGAACTTGGCCGCGACCGC 420
Qy 421 TCATGAGCTGGTGTGCGGGCTTCTGAGAGGCAAGAGCAAGCGGCTGCGGGC 480
Db 421 TCATGAGCTGGTGTGCGGGCTTCTGAGAGGCAAGAGCAAGCGGCTGCGGGC 480
Qy 481 GCAAGCTCATCCGCTACCGCAACCTGGGCAACGCTCATCTGCGGACGCTGACAC 540
Db 481 GCAAGCTCATCCGCTACCGCAACCTGGGCAACGCTCATCTGCGGACGCTGACAC 540
Qy 541 CAGTCTAACAGGCGCTTCCCAAGCGCCAGCACTGCTGCAAGAGCTTTATGACTCGG 600
Db 541 CAGTCTAACAGGCGCTTCCCAAGCGCCAGCACTGCTGCAAGAGCTTTATGACTCGG 600
Qy 601 CAGAAACACAGAGCTTGAAGAACTGAGCTTACCAACATGTTCTGGGTGACCTGG 660
Db 601 CAGAAACACAGAGCTTGAAGAACTGAGCTTACCAACATGTTCTGGGTGACCTGG 660
Qy 661 TGTGTTTGGCCACCTGTCATGAAAGGCTGGCTTGAAGAGTGGAGCTTATCC 720
Db 661 TGTGTTTGGCCACCTGTCATGAAAGGCTGGCTTGAAGAGTGGAGCTTATCC 720
Qy 721 TGTCTCAGAGCTGCTGAACAGATGAACCTTGGCGTCACTGATGAGCACTGTATG 780
Db 721 TGTCTCAGAGCTGCTGAACAGATGAACCTTGGCGTCACTGATGAGCACTGTATG 780
Qy 781 CCTACGACTGATTAATATCCACTGCTGTATACAGAGTGTGACTGTGGCGGTGACA 840
Db 781 CCTACGACTGATTAATATCCACTGCTGTATACAGAGTGTGACTGTGGCGGTGACA 840
Qy 841 GCTTCTTCCGACCTTGTCTAATGTTGGGGGCGAGTTTCTGAACCCAGCAAGCTTAC 900
Db 841 GCTTCTTCCGACCTTGTCTAATGTTGGGGGCGAGTTTCTGAACCCAGCAAGCTTAC 900
Qy 901 GCGATGAGCTGAGCTGCTGTTGGCCGTTCACTTCAAGTTCTTCTTAAGTTG 960
Db 901 GCGATGAGCTGAGCTGCTGTTGGCCGTTCACTTCAAGTTCTTCTTAAGTTG 960
Qy 961 GCTGGCTGAAGGTGGCAGAGCACTCAACCCCTTGAAGAGATGATGATTTTG 1020
Db 961 GCTGGCTGAAGGTGGCAGAGCACTCAACCCCTTGAAGAGATGATGATTTTG 1020
Qy 1021 AGACCAACTGATTTGAGACAGAAATTTGAGAGTGTCCCTGGCTGTGAGATGAGATGC 1080
Db 1021 AGACCAACTGATTTGAGACAGAAATTTGAGAGTGTCCCTGGCTGTGAGATGAGATGC 1080
Qy 1081 ACCAGAGCTGCTCGATGAGAGCCGAGCATGTACTGGAATTAAGCCGAGCCAGCCCC 1140
Db 1081 ACCAGAGCTGCTCGATGAGAGCCGAGCATGTACTGGAATTAAGCCGAGCCAGCCCC 1140

Db 1081 ACCAGAGCTGCTCGATGAGAGCCGAGCATGTACTGGAATTAAGCCGAGCCAGCCCC 1140
Qy 1141 CCTACAGAGCTTCCGCCCAAGTTCCGTGAGAGCTCTTTATGAGCTTCCACCTTAACA 1200
Db 1141 CCTACAGAGCTTCCGCCCAAGTTCCGTGAGAGCTCTTTATGAGCTTCCACCTTAACA 1200
Qy 1201 TCAGCTGAACAAAGAGAGATGAGTTCAGGCCCAATCAAGAGAGAGAGATGCTC 1260
Db 1201 TCAGCTGAACAAAGAGAGATGAGTTCAGGCCCAATCAAGAGAGAGAGATGCTC 1260
Qy 1261 AGCTGGCATCATTTGGCCGCTTCTGAGGCTGAGAGTCCATGATCACTCTCCAGGG 1320
Db 1261 AGCTGGCATCATTTGGCCGCTTCTGAGGCTGAGAGTCCATGATCACTCTCCAGGG 1320
Qy 1321 CAAACTCAAGAGCAAACTGCTGGCCCAAGAGGAAATCCTTCTTCCAGAGAGGCTGC 1380
Db 1321 CAAACTCAAGAGCAAACTGCTGGCCCAAGAGGAAATCCTTCTTCCAGAGAGGCTGC 1380
Qy 1381 CCAAAAACCAAGGCGCCCAACAGAGGTTAGGGGCGAGAGAACAAAGCCCTGGA 1440
Db 1381 CCAAAAACCAAGGCGCCCAACAGAGGTTAGGGGCGAGAGAACAAAGCCCTGGA 1440
Qy 1441 AGCTTAAGCTGTGAGAGCTTCAAGTCTGGCCCACTGATCAGAGGCGCTACTGA 1500
Db 1441 AGCTTAAGCTGTGAGAGCTTCAAGTCTGGCCCACTGATCAGAGGCGCTACTGA 1500
Qy 1501 GTGCCCCAGAGAGGCCCCCTCAGGCCCATCCATGTTCTTCCCTTGAACCATCAGCGC 1560
Db 1501 GTGCCCCAGAGAGGCCCCCTCAGGCCCATCCATGTTCTTCCCTTGAACCATCAGCGC 1560
Qy 1561 CGTCAAGCTTCACTAGTCAAGAGATGACCAAGAGCAAAAGCTTAAAGCTGTA 1620
Db 1561 CGTCAAGCTTCACTAGTCAAGAGATGACCAAGAGCAAAAGCTTAAAGCTGTA 1620
Qy 1621 GTTCTGGGGCAAGAAAGTTTGAATGCTCTCAGAGAGCATGGGCTTGTATGAGC 1680
Db 1621 GTTCTGGGGCAAGAAAGTTTGAATGCTCTCAGAGAGCATGGGCTTGTATGAGC 1680
Qy 1681 ACCCAGAGATCTCAAGTGAAGAGAGAACTGTGAGTTTAACTGACGGTATGCGAG 1740
Db 1681 ACCCAGAGATCTCAAGTGAAGAGAGAACTGTGAGTTTAACTGACGGTATGCGAG 1740
Qy 1741 AGATCCCGCAAAATCACTCAAGAACTTTGGAACAATCAACAACATACACACTA 1800
Db 1741 AGATCCCGCAAAATCACTCAAGAACTTTGGAACAATCAACAACATACACACTA 1800
Qy 1801 CACTCAAAATCACTGATCTTATTTGGGCTTGGAAAACAGGATGAAGCAATTCCT 1860
Db 1801 CACTCAAAATCACTGATCTTATTTGGGCTTGGAAAACAGGATGAAGCAATTCCT 1860
Qy 1861 AACCTGCTTCTTAATGGGGAGTCTTGGCAGCAGGCTCTCACTGCTGTATACACAGA 1920
Db 1861 AACCTGCTTCTTAATGGGGAGTCTTGGCAGCAGGCTCTCACTGCTGTATACACAGA 1920
Qy 1921 GGAACATGATCCAGTCAAGCATACAGCTGTCCAACATGAAAGAGCTGTCTCAACAG 1980
Db 1921 GGAACATGATCCAGTCAAGCATACAGCTGTCCAACATGAAAGAGCTGTCTCAACAG 1980
Qy 1981 CCTGAATCAAAAGTTAGCTTATATGATTAATAATCCAGACTTACGCTTTATATCC 2040
Db 1981 CCTGAATCAAAAGTTAGCTTATATGATTAATAATAATCCAGACTTATATATGCC 2040
Qy 2041 TTTTATTAATAAAACCTGTAAGCTGATGAACTGAAACATTTAATCTAGACTC 2100
Db 2041 TTTTATTAATAAAACCTGTAAGCTGATGAACTGAAACATTTAATCTAGACTC 2100
Qy 2101 TGGATTCAAGAGTGGGAAACCTTATGTTCTATCTGAATCAAGACGACACTTAATAT 2160
Db 2101 TGGATTCAAGAGTGGGAAACCTTATGTTCTATCTGAATCAAGACGACACTTAATAT 2160
Qy 2161 ACTGCCCAAACTAATGAGTTTAAATTAATAAATCAAAATCTGTTAAAAA 2220
Db 2161 ACTGCCCAAACTAATGAGTTTAAATTAATAAATCAAAATCTGTTAAAAA 2220

QY	2221	AAAAAAAA	2229	
Db	2221	AAAAAAAA	2229	
RESULT 2	AF057169			
LOCUS	AF057169			
DEFINITION	2210 bp	mRNA	linear	PRI 17-OCT-1998
ACCESSION	AF057169			
VERSION	AF057169			
KEYWORDS	AF057169.1	GI:3335158		
SOURCE				
ORGANISM	Homio sapiens	(human)		
REFERENCE	1	Homio sapiens		
AUTHORS	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1	(bases 1 to 2210)		
JOURNAL	1	Petrukhin, K., Koist, M.J., Bakall, B., Li, W., Xie, G., Marknell, T., Sandgren, O., Foreman, K., Holmren, G., Andreasson, S., Vujic, M., Metzker, M.L., Caskey, C.T. and Wadelius, C.		
MEDLINE	98324772			
PUBMED	9663395			
REFERENCE	2	(bases 1 to 2210)		
AUTHORS	1	Petrukhin, K.		
TITLE	1	Direct Submision		
JOURNAL	1	Submitted (03-APR-1998) Human Genetics, Merck Research Laboratories, West Point, PA 19486, USA		
FEATURES				
source	1..2210	location/Qualifiers		
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/chromosome="11"			
	/map="11q13"			
	1..2210			
	/gene="VMD2"			
	105..1862			
	/gene="VMD2"			
	/note="Best's macular dystrophy gene; BMD; alternatively spliced isoform"			
	/codon_start=1			
	/product="dystrophin"			
	/protein_id="AAC64343.1"			
	/db_xref="GI:3335159"			
	translation="MTITYSOVANARLGSFRLLCRGSITYKLLYGEPLIFLLCYVQIIRIPRYALTEEOQIMPEKTLICVDSYICLIPISFVIGPVTLVTRMKNQVLPYEDRLMSLVSGFVEGDEQGRILRLRLITRYALNGVLLIRSVTAHYRFPBAQHVOA			
	GTPCTAEHKOLEKLSLPNNMFVPPVWPAFLSNKAWLGRJRDPLLOSILNENVTLLTFCGLLAVYDVISPLVYTVVTAVASFPFLTCLVGRFLNPPAKVPEHELDLVVPEVTFLOGLFFVYGM/LKVAEQLINPEGEDDDDEFTWMLIVDRMLQVSLAVDEMODLPRMEBETDMMWPKPEPOPPTAASQFPRASFGSTFNIISLKEEMEROPNDEDEDAHAGIGRFLGLOSHDHPRAANSPTKLMPKRESLTHGLPKNHKLAQONRGCDENKAWMLKAA			
	DAKSAEPYQRPRTYSABQTPLSPTSPMPFPLPSPAPSKLHAYTGIDTDKSLKYSSGAKSSELSSESDEGALMEHPVSVQRKTVMEFNLDPMPPIPNHLKEPLEGSPYINHTT			
	LKDHPDPYALNENRDEAHS"			
ORIGIN				
Query Match	99.1%	Score 2208.4;	DB 9;	Length 2210;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2209;	Conservative	0;	Mismatches	1; Indels 0; Gaps 0.
QY	1	CAGGAGTCCCAACGACCTAGTCGCAAGACCTTCTGTGGAGATCATCGAACCACTGGAA	60	
Db	1	CAGGAGTCCCAACGACCTAGTCGCAAGACCTTCTGTGGAGATCATCGAACCACTGGAA	60	
QY	61	CCCCACCTGACCCGAAGCCCACTGCTGCGACGCCCACTGCTGGCCATGACCATCTTACA	120	
Db	61	CCCCACCTGACCCGAAGCCCACTGCTGCGACGCCCACTGCTGGCCATGACCATCTTACA	120	

QY	121	CAAGCAGATGGCTAAATGCGCGCTTAAGGCTCCTTCTCCGCGCTGCGCTGCTGCGCGG	180
Db	121	CAGCGAAGTGGCTAAATGCGCGCTTAAGGCTCCTTCTCCGCGCTGCGCTGCTGCGCGG	180
QY	181	GCAGCATCTACAAAGCTGCTATATGCGAGATTCTTAATCTTCTCGTCTGCTACTACATCA	240
Db	181	GCAGCATCTACAAAGCTGCTATATGCGAGATTCTTAATCTTCTCGTCTGCTACTACATCA	240
QY	241	TCCGCTTTATTTATATGCGTGGCCCTCACGGAAGAACAAAGCTGATGTTTGAGAACTGA	300
Db	241	TCCGCTTTATTTATATGCGTGGCCCTCACGGAAGAACAAAGCTGATGTTTGAGAACTGA	300
QY	301	CTCTGTATTTGCGACAGCTACATCCAGCTCATCCCAATTTCTCTCGTGGCTGCGCTTACG	360
Db	301	CTCTGTATTTGCGACAGCTACATCCAGCTCATCCCAATTTCTCTCGTGGCTGCGCTTACG	360
QY	361	TGACGCTGTCGTGACCCCGCTGGTGAACCAAGTACGAAGAACTGCGCTGCGACCGCTC	420
Db	361	TGACCGCTGTCGTGACCCCGCTGGTGAACCAAGTACGAAGAACTGCGCTGCGACCGCTC	420
QY	421	TCATGAGCTGTGTGTCGGGCTTCGTGGAAGGACGAAGAGAGGACGCGGCTGCTGCGG	480
Db	421	TCATGAGCTGTGTGTCGGGCTTCGTGGAAGGACGAAGAGAGGACGCGGCTGCTGCGG	480
QY	481	GCACGCTCATCCGCTACACGCCAACTGCGGCAACGTGCTCATCTCTGCGCAGCGTACGACCG	540
Db	481	GCACGCTCATCCGCTACACGCCAACTGCGGCAACGTGCTCATCTCTGCGCAGCGTACGACCG	540
QY	541	CAGTTTACAAGCGCTTCCCCAGCGGCCACAGCACTGTCGACAGAGCTTTATGATCTCGG	600
Db	541	CAGTTTACAAGCGCTTCCCCAGCGGCCACAGCACTGTCGACAGAGCTTTATGATCTCGG	600
QY	601	CAGAACACAAGCATTTGAGAAACTGAGGCTACACACAATGTTCTTGCGTCCCTCGG	660
Db	601	CAGAACACAAGCATTTGAGAAACTGAGGCTACACACAATGTTCTTGCGTCCCTCGG	660
QY	661	TGTGTTTGCCAACTGTCAATGAAGCGGTGGCTTGAGGTCGAATCCGGGACCTTATCC	720
Db	661	TGTGTTTGCCAACTGTCAATGAAGCGGTGGCTTGAGGTCGAATCCGGGACCTTATCC	720
QY	721	TGCTCCAGAGCTGTGTAACGAGATGAACAACCTTGCGTACTGACAGTGTGACACCTGTATG	780
Db	721	TGCTCCAGAGCTGTGTAACGAGATGAACAACCTTGCGTACTGACAGTGTGACACCTGTATG	780
QY	781	CCTACGACTGGAATTAGTATCCCACTGATGTATACAAGGTGTGACTGTGCGCGGTATACA	840
Db	781	CCTACGACTGGAATTAGTATCCCACTGATGTATACAAGGTGTGACTGTGCGCGGTATACA	840
QY	841	GCTTCTTCTGACTTTGTCTAGTTGGGCGGCACTTTCTGAACCCAGCCAAAGGCTTACCTTG	900
Db	841	GCTTCTTCTGACTTTGTCTAGTTGGGCGGCACTTTCTGAACCCAGCCAAAGGCTTACCTTG	900
QY	901	GCATGAGCTGGAACCTGCTGTCGCCGCTTCAACGTTCTGTGAGTTCCTTCTTATGTTG	960
Db	901	GCATGAGCTGGAACCTGCTGTCGCCGCTTCAACGTTCTGTGAGTTCCTTCTTATGTTG	960
QY	961	GCTGGCTAAGGTGGCAGACGACTCATCAACCCCTTTGAGAGGAGATGATGATTTTG	1020
Db	961	GCTGGCTAAGGTGGCAGACGACTCATCAACCCCTTTGAGAGGAGATGATGATTTTG	1020
QY	1021	AGACCAATGGAATTTGTGACAGGAATTTTGACAGTGTCCCTGTTGGCTGTGGAATGAGATGC	1080
Db	1021	AGACCAATGGAATTTGTGACAGGAATTTTGACAGTGTCCCTGTTGGCTGTGGAATGAGATGC	1080
QY	1081	ACAGAGACTGTGCTGATGAGACCGGACATGTACTGGAATTAAGCCCGACCAAGCCCC	1140
Db	1081	ACAGAGACTGTGCTGATGAGACCGGACATGTACTGGAATTAAGCCCGACCAAGCCCC	1140
QY	1141	CCTACAGAGCTGTGCTTCCGCCCACTGTTCCGTGAGCTCTCTTATGAGGCTCCACTTCAACA	1200
Db	1141	CCTACAGAGCTGTGCTTCCGCCCACTGTTCCGTGAGCTCTCTTATGAGGCTCCACTTCAACA	1200

QY 1201 TCAGCCTGAACAAAGAGAGATGATGATCCAGCCCAATCATGAGAGAGAGAGATGCTC 1260
Db 1201 TCAGCCTGAACAAAGAGAGATGATGATCCAGCCCAATCATGAGAGAGAGAGATGCTC 1260
QY 1261 ACCGTGGCATCATTTGGCCGCTTCTTGAAGGCTGCAATGATGATCAATCCTCCAGAG 1320
Db 1261 ACCGTGGCATCATTTGGCCGCTTCTTGAAGGCTGCAATGATGATCAATCCTCCAGAG 1320
QY 1321 CAAACTCAAGAGCAAACTACTGTGGCCCAAGAGAGAGATCCCTTCCCAAGAGAGAGCTC 1380
Db 1321 CAAACTCAAGAGCAAACTACTGTGGCCCAAGAGAGAGATCCCTTCCCAAGAGAGAGCTC 1380
QY 1381 CCAAAAACCAAG 1440
Db 1381 CCAAAAACCAAG 1440
QY 1441 AGCTTAAGGCTGAG 1500
Db 1441 AGCTTAAGGCTGAG 1500
QY 1501 GTGGCCCAAG 1560
Db 1501 GTGGCCCAAG 1560
QY 1561 CGTCAAAAGCTTCAAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 CGTCAAAAGCTTCAAGATGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 GTTCTGGGGGCAAG 1680
Db 1621 GTTCTGGGGGCAAG 1680
QY 1681 ACCCAAGAGATATCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1681 ACCCAAGAGATATCTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 1741 AGATCCCGGAAATCAACCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 1741 AGATCCCGGAAATCAACCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1801 CACTCAAAAGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1801 CACTCAAAAGATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
QY 1861 AACCTGCTTCCAAAG 1920
Db 1861 AACCTGCTTCCAAAG 1920
QY 1921 GAGACATGATCAAGTCAAG 1980
Db 1921 GAGACATGATCAAGTCAAG 1980
QY 1981 CCTGAATCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db 1981 CCTGAATCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
QY 2041 TTTTATTCATTAATAAAG 2100
Db 2041 TTTTATTCATTAATAAAG 2100
QY 2101 TGGATTCAGAGTGGGAAACCTTATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2101 TGGATTCAGAGTGGGAAACCTTATGATGATGATGATGATGATGATGATGATGATGAT 2160
QY 2161 ACTGCCCAAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2210
Db 2161 ACTGCCCAAACTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2210

RESULT 3
AF073501 2171 bp mRNA linear PRI 02-SEP-1998
LOCUS Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA,
DEFINITION

complete cds.
AF073501
AF073501.1 GI:3511241
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Stoehr, H., Marguard, A., Rivera, A., Cooper, P. R., Nowak, N. J.,
Shows, T. B., Gerhard, D. S. and Weber, B. H.,
A gene map of the Best's vitelliform macular dystrophy region in
chromosome 11q12-q13.1
Genome Res. 8 (1), 48-56 (1998)
98112782
MEDLINE
PUBMED
9445487
2 (bases 1 to 2171)
Marguard, A., Stoehr, H., Passmore, L., Kraemer, F., Rivera, A. and
Weber, B. H.,
Direct Submission
Submitted (22-JUN-1998) Human Genetics, University, Biozentrum, Am
Hudland, Wuerzburg 97074, Germany
location/Qualifiers
1. .2171
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12-q13.1"
/tissue_type="retinal pigment epithelium"
1. .2171
/gene="VMD2"
89. .1846
/gene="VMD2"
/note="Best disease"
/codon_start=1
/product="vitelliform macular dystrophy protein"
/protein_id="AAC33766.1"
/db_xref="GI:3511242"
/translation="MTITVTSQVNAARLGSRSRLLCMRGSIYKLLYSEFLIFLCY
ITRFLRYALTEBOQIMBERKTLTYCDSTYQLPISFVIGFVTVLVTVMMQYENLP
PRLMSLVSGFVEGDEBQRLRLRLRYANLGNVLIRSVTAHYKSPSAOHLYVA
GMTPAEHKOELKSLPHNMFVPMVPMFANLSMKAMLGRIIDPILLOSILNEMNTR
TQCHLYAYDWISIPLYTVQVTVAVSFEFLICVLRQFLNPAPAKYPGHELDLVVPE
TELOPFYVGMKVAEOLINPEGDEDDFEINMIVDRNLQVSLAVDEMDLPRME
DVMNKPBPQPYTAASQFRRASPMSTPMNLSINKEMERQPNQDEEDHAAGTIGR
FLGLSHDHHPRAHSRKLMLMPKRESLHSGLPKRNHAKQNVRGQSDNKAAMLKAV
DAFKSAPIYORPGTYSAPQTPUSPTPMFPLEPSPAPSKLHVTGIDTKSLKTVSSG
AKGSFELSDGALMEHPVSVQVARKTVEENLTDMEIPENHLKEPLEOSPTIHTT
IKDHMDPYWALENRDEAHS"

ORIGIN

Query Match 97.2%; Score 2166.2; DB 9; Length 2171;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 17 CCTAGTGGCAGACCTTCTGTGGATCATCGAGACCACTGGAACCCCACTGACCCCAAG 76
Db 1 CCTAGTGGCAGACCTTCTGTGGATCATCGAGACCACTGGAACCCCACTGACCCCAAG 60
QY 77 CCCACTGCTGACAGCCCACTGCTGGCCATGACATCATTTACACAAAGAGGCTAA 136
Db 61 CCCACTGCTGACAGCCCACTGCTGGCCATGACATCATTTACACAAAGAGGCTAA 120
QY 137 TGCCCGCTTAAAGTCTCTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 196
Db 121 TGCCCGCTTAAAGTCTCTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 197 GCTATATGGGAGATTTTAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 256
Db 181 GCTATATGGGAGATTTTAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 257 GCTGCGCCCTACGGAAGAACAGCTGATGTTTGAAGAACTGACTGTGATTTGGCAGAG 316

|||||
Db 241 GCTGGCCTCA CGAAGAACAA CAGCTGATGTTGAGAAATGACTGTATTCGACAG 300
|||
Qy 317 CTATATCAGAGTATCCCATATTTCTTCGTCTGTGGCTTCACTGACGCTGTGTGAC 376
|||
Db 301 CTATATCAGAGTATCCCATATTTCTTCGTCTGTGGCTTCACTGACGCTGTGTGAC 360
|||
Qy 377 CCGGTGTGGAACAGTACAGAACTGCGTGGCCCGACCGCTCACTGAGCTGTGTG 436
|||
Db 361 CCGGTGTGGAACAGTACAGAACTGCGTGGCCCGACCGCTCACTGAGCTGTGTG 420
|||
Qy 437 GGGCTTGTGGAAGCAAGAGCAAGCAAGCGCGCTGTGGGCGCAAGCTCATCCGCTA 496
|||
Db 421 GGGCTTGTGGAAGCAAGAGCAAGCAAGCGCGCTGTGGGCGCAAGCTCATCCGCTA 480
|||
Qy 497 CGCCAACTGGGCAAGCTGTCTATCTTCGCGAGGCTGACGACCCGCACTCAAGCGCTT 556
|||
Db 481 CGCCAACTGGGCAAGCTGTCTATCTTCGCGAGGCTGACGACCCGCACTCAAGCGCTT 540
|||
Qy 557 CCCAGAGCCCAAGCAAGCTGTGCAAGCAAGCTTATGACTCCGCGAAGAACAGCAAGT 616
|||
Db 541 CCCAGAGCCCAAGCAAGCTGTGCAAGCAAGCTTATGACTCCGCGAAGAACAGCAAGT 600
|||
Qy 617 GGAAGAACTGAGCTTACCAACAATGTTCTGGGTGCGCTGGGTGTGCTTTCGCAACT 676
|||
Db 601 GGAAGAACTGAGCTTACCAACAATGTTCTGGGTGCGCTGGGTGTGCTTTCGCAACT 660
|||
Qy 677 GTCAATGAAAGCGTGGCTGTGAGGTGCAATCCGGAACCTTATCTCTGCTCCAGAGCTGCT 736
|||
Db 661 GTCAATGAAAGCGTGGCTGTGAGGTGCAATCCGGAACCTTATCTCTGCTCCAGAGCTGCT 720
|||
Qy 737 GAAAGAGTGAACAACCTTGGGCTGACTCAGTGTGGAACCTGTATGCTCAAGCTGATTAAG 796
|||
Db 721 GAAAGAGTGAACAACCTTGGGCTGACTCAGTGTGGAACCTGTATGCTCAAGCTGATTAAG 780
|||
Qy 797 TATCCACTGTGTATACACAGTGTGACTGTGGCGGTGTACAGCTTCTTCTGACTTG 856
|||
Db 781 TATCCACTGTGTATACACAGTGTGACTGTGGCGGTGTACAGCTTCTTCTGACTTG 840
|||
Qy 857 TCTAGTGTGGGCGGCAAGTTCTTGAAACCAAGCCCAAGGCTTACCTGCGCATGAGCTGACCT 916
|||
Db 841 TCTAGTGTGGGCGGCAAGTTCTTGAAACCAAGCCCAAGGCTTACCTGCGCATGAGCTGACCT 900
|||
Qy 917 CGTGTGGCGGCTTCAAGTCTTCAAGTCTTCTGATGTTGGCTGGAGAGGTGGC 976
|||
Db 901 CGTGTGGCGGCTTCAAGTCTTCAAGTCTTCTGATGTTGGCTGGAGAGGTGGC 960
|||
Qy 977 AGAGCACTCATCAACCCCTTGGAGAGATGATGATTTTGACCAACTGGAATTTGT 1036
|||
Db 961 AGAGCACTCATCAACCCCTTGGAGAGATGATGATTTTGACCAACTGGAATTTGT 1020
|||
Qy 1037 CGACAGAAATTTGCAAGGTGTCCTGTGGCTGTGGATGAGATGACCAAGCACTGCTCG 1096
|||
Db 1021 CGACAGAAATTTGCAAGGTGTCCTGTGGCTGTGGATGAGATGACCAAGCACTGCTCG 1080
|||
Qy 1097 GATGGAAGCCGAGCAATGTAAGTGAATTAAGCCGAGCCCAAGCCCTTCAAGCTGCTC 1156
|||
Db 1081 GATGGAAGCCGAGCAATGTAAGTGAATTAAGCCGAGCCCAAGCCCTTCAAGCTGCTC 1140
|||
Qy 1157 CGCCCAATTCGCTGAGAGCTCTTATGAGGCTCCACCTTCAACATCAAGCTGAACAAAGA 1216
|||
Db 1141 CGCCCAATTCGCTGAGAGCTCTTATGAGGCTCCACCTTCAACATCAAGCTGAACAAAGA 1200
|||
Qy 1217 GGAAGATGAGATTTCAAGCCCAATCAAGAGAGCAAGAGAGATCTCAAGCTGCAATTTGG 1276
|||
Db 1201 GGAAGATGAGATTTCAAGCCCAATCAAGAGAGCAAGAGAGATCTCAAGCTGCAATTTGG 1260
|||
Qy 1277 CCGCTTCTTAGGCGTGAAGTCCCAATCAAGATCAACATCTTCCAGGGCAACTGAAGCAACA 1336
|||
Db 1261 CCGCTTCTTAGGCGTGAAGTCCCAATCAAGATCAACATCTTCCAGGGCAACTGAAGCAACA 1320
|||
Qy 1337 ACTACTGTGGCCAAAGAGGAAATCCCTTCTCAAGAGGCTGTGCCAAAACCAAGAGC 1396
|||

Db 1321 ACTACTGTGGCCAAAGAGGAAATCCCTTCTCAAGAGGCGCTGTGCCAAAACCAAGGC 1380
|||
Qy 1397 AGCCAAACAGAACGTTAGGGGCGCAGAAAGCAACAAAGGCTGAAAGCTTAAGGCTGTGGA 1456
|||
Db 1381 AGCCAAACAGAACGTTAGGGGCGCAGAAAGCAACAAAGGCTGAAAGCTTAAGGCTGTGGA 1440
|||
Qy 1457 CGCCTTCAAGTCTGAGCCCACTGTATCAAGAGCCAGGCTACTACAGTSCCCCAAGAGGCC 1516
|||
Db 1441 CGCCTTCAAGTCTGAGCCCACTGTATCAAGAGCCAGGCTACTACAGTSCCCCAAGAGACC 1500
|||
Qy 1517 CTTGAGCCCACTTCCCATGTTCTTCCCTTGAAGCAATCAAGCCGCTTCAAGCTTTCACAG 1576
|||
Db 1501 CTTGAGCCCACTTCCCATGTTCTTCCCTTGAAGCAATCAAGCCGCTTCAAGCTTTCACAG 1560
|||
Qy 1577 TGTCAAGGCAATGAGCAACAAAGCAAAAGCTTAAAGCTGTGAGTTCTGTGGGCGCAAGAA 1636
|||
Db 1561 TGTCAAGGCAATGAGCAACAAAGCAAAAGCTTAAAGCTGTGAGTTCTGTGGGCGCAAGAA 1620
|||
Qy 1637 AAGTTTGAATTTGCTCTCAAGAGCGATGGGCGCTTGTATGAGAGCACCCAGAAATATCTCA 1696
|||
Db 1621 AAGTTTGAATTTGCTCTCAAGAGCGATGGGCGCTTGTATGAGAGCACCCAGAAATATCTCA 1680
|||
Qy 1697 AGTGAAGAGAAACTGTGAGTTTAACTGACGGAATATGCGAGATATCCCGAAATCA 1756
|||
Db 1681 AGTGAAGAGAAACTGTGAGTTTAACTGACGGAATATGCGAGATATCCCGAAATCA 1740
|||
Qy 1757 CTTCAAGAAACCTTTGGAACAATTCAGAACCAATCACTACACTCAAAAGATCAAT 1816
|||
Db 1741 CTTCAAGAAACCTTTGGAACAATTCAGAACCAATCACTACACTCAAAAGATCAAT 1800
|||
Qy 1817 GSAATCTTATTTGGGCTTGGAAAAACAGGATGAAGCAATTTCTTAACTGCTTCTTATG 1876
|||
Db 1801 GSAATCTTATTTGGGCTTGGAAAAACAGGATGAAGCAATTTCTTAACTGCTTCTTATG 1860
|||
Qy 1877 GGGATGCTTGGCCAGCCAGAGTCTCACTGTGTGTATACACAGCAAGCACTGATCCAGTGC 1936
|||
Db 1861 GGGATGCTTGGCCAGCCAGAGTCTCACTGTGTGTATACACAGCAAGCACTGATCCAGTGC 1920
|||
Qy 1937 ACAGCCATACAGCTGTCCACATGAAAGAGTGTCTTCAACAAGCCGAAATCAATGGTT 1996
|||
Db 1921 ACAGCCATACAGCTGTCCACATGAAAGAGTGTCTTCAACAAGCCGAAATCAATGGTT 1980
|||
Qy 1997 AGCTTAATAGATAAAATCCCAAGTACTTCAAGCTTAAATGCTTATTTATTCATTAANAAC 2056
|||
Db 1981 AGCTTAATAGATAAAATCCCAAGTACTTCAAGCTTAAATGCTTATTTATTCATTAANAAC 2040
|||
Qy 2057 TGTGAAGCTTACAGTGAACCAATTTGAACAATTTAACTCAAGCTGTGATTCAGAGTGGG 2116
|||
Db 2041 TGTGAAGCTTACAGTGAACCAATTTGAACAATTTAACTCAAGCTGTGATTCAGAGTGGG 2100
|||
Qy 2117 AACCTTAAGTTCTATCTGAATCCAAAGCAGCCACCTTAGTAACTGCCCCAACTAATG 2176
|||
Db 2101 AACCTTAAGTTCTATCTGAATCCAAAGCAGCCACCTTAGTAACTGCCCCAACTAATG 2160
|||
Qy 2177 AGTTTAATAA 2187
|||
Db 2161 AGTTTAATAA 2171
|||

RESULT 4
C0729633 2170 bp DNA linear PAT 03-FEB-2004
LOCUS C0729633
DEFINITION Sequence 15567 from Patent WO02068579.
ACCESSION C0729633
VERSION C0729633.1 GI:42301252
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
Kits, such as nucleic acid arrays, comprising a majority of

humanexons or transcripts, for detecting expression and other uses thereof
Patent: WO 02068579-A 15567 06-SEP-2002;
DE 10034232 (1999) (1999)

FEATURES	Location/Qualifiers
source	1. .2170

```

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

Query Match	96.7%;	Score 2155.8;	DB 6;	Length 2170;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 2168; Conservative	0;	Mismatches 2;	Indels 1.	Coverage 1

QY	17	CGTAGTGGCCAGACCTTCTGTGGATCATCGAACCCACTGGAAACCCACTGGACCCAG	76
Db	1	CTTAGTGGCCAGACCTTCTGTGGATCATCGAACCCACTGGAAACCCACTGGACCCAG	60
QY	77	CCCACTGCTGCAGCCCACTGCCTGGCCATGACCACTACCTTATACAAAGCCAACTGGCTTA	136
Db	61	CCCACTGCTGCAGCCCACTGCCTGGCCATGACCACTACCTTATACAAAGCCAAAG-AGCTAA	119
QY	137	TGCCCGTTAGGCTCCTTCTTCCCGCCTGCTGCTGTCCTGGCGGGGAGCATCTAAGCT	196
Db	120	TGCCCGTTAGGCTCCTTCTTCCCGCCTGCTGCTGTCCTGGCGGGGAGCATCTAAGCT	179
QY	197	GCTAATAGCGAGTCTTAACTCTTCCGCTGCTGCTATACATCAATCCGTTATTATTAAG	256
Db	180	GCTAATAGCGAGTCTTAACTCTTCCGCTGCTGCTATACATCAATCCGTTATTATTAAG	239
QY	257	GCAGGCCCTCAGCGAAGAAACAACGCTATGTTTGAAGAACTGACTGTGTAATGGACAG	316
Db	240	GCAGGCCCTCAGCGAAGAAACAACGCTATGTTTGAAGAACTGACTGTGTAATGGACAG	299
QY	317	CTACATCGACTCATCCCATTTCTTCTGTGCTGGACCTTAACTGAGAGCTGTCGTGAC	376
Db	300	CTACATCGACTCATCCCATTTCTTCTGTGTCGTGGCTTAACTGAGAGCTGTCGTGAC	359
QY	377	CCGCTGGTGAACCCAGTACAGAACACTTGCCGTGGCCGACCGCTCATGACCTGATGC	436
Db	360	CCGCTGGTGAACCCAGTACAGAACACTTGCCGTGGCCGACCGCTCATGAGCTGATGC	419
QY	437	GGGCTTTCGTGGAAGGCAAGACAGACAGCCGCGCTGTGAGGGCAACGCTATCCGCTA	496
Db	420	GGGCTTTCGTGGAAGGCAAGACAGACAGCCGCGCTGTGAGGGCAACGCTATCCGCTA	479
QY	497	CGCCAAACCTGGGCAACCGTGTCTATCTCTGCGACGCTGACACCGCAGTCTTAAAGCGCTT	556
Db	480	CGCCAAACCTGGGCAACCGTGTCTATCTCTGCGACGCTGACACCGCAGTCTTAAAGCGCTT	539
QY	557	CCCCAGCGCCAGACACTGTGTGCAAGAGCTTATGATCCCGGCAAGAACAAAGAGATT	616
Db	540	CCCCAGCGCCAGACACTGTGTGCAAGAGCTTATGATCCCGGCAAGAACAAAGAGATT	599
QY	617	GGAGAAATGAGCCTACCAACAACATGTTCTGGGTGCTCCCTGGAGTGTGTTGCCAACT	676
Db	600	GGAGAAATGAGCCTACCAACAACAATGTTCTGGGTGCTCCCTGGAGTGTGTTGCCAACT	659
QY	677	GTCATGAAAGCGTGGCTTGGAGGTTCGAATCCGAGAACCTTATCTGTCTCAAGAGCTGCT	736
Db	660	GTCATGAAAGCGTGGCTTGGAGGTTCGAATCCGAGAACCTTATCTGTCTCAAGAGCTGCT	719
QY	737	GAAAGAGATGAAACCTTGGGTATCTAGTGTGACACTGTATACCTTACGACTGATTAAG	796
Db	720	GAAAGAGATGAAACCTTGGGTATCTAGTGTGACACTGTATACCTTACGACTGATTAAG	779
QY	797	TATCCCACTGTGTATACACAGGTGTGACTGTGGAGGTGATCAGCTTCTCTGACCTTG	856
Db	780	TATCCCACTGTGTATACACAGGTGTGACTGTGGAGGTGATCAGCTTCTCTGACCTTG	839
QY	857	TCTAGTGGGCGCAGTTTCTGAACCCAAACCAAGGCTTACCTGGCAATGACTGACCT	916

Db	840	TTTAATTGAGGCGAGATTTC	TGAACCCAGCAAGGCGCTAC	CTCGGCGCATGAGTCGACCT	899
QY	917	CGTTGTGCGCCCTCTTCA	CGTTCCTGACAGTTCCTT	CTTCTATGTTGGGTGCGTAAAGTGCG	976
Db	900	CGTTGTGCGCGCTTTC	TCAAGTTCC	TGCAAGTTCTTCTTCTATGTTGGGTGCGTAAAGTGCG	959
QY	977	AGAGAGCTCATCA	CCCCCTTTGGAGAGATGAT	ATGATTTTGGAGCAACCTGATATGT	1036
Db	960	AGAGAGCTCATCA	ACCCCTTTGGAGAGATGAT	ATGATTTTGGAGCAACCTGATATGT	1019
QY	1037	CGACAGAAATTTGACAGGTG	CCCTGTTGGGTGGA	TGAGATGACACAGACTTCCCTCG	10986
Db	1020	GCAAGAGAAATTTGAGAGT	GTCCCTGTTGGGTGGA	TGAGATGACACAGACTTCCCTCG	10799
QY	1097	GATGAGACCCGACATGTA	CTGGAATTAAGCCCGAGCCACAGCCCCCTTACAGAGCTGCTTC	1156	
Db	1080	GATGAGACCCGACATGTA	CTGGAATTAAGCCCGAGCCACAGCCCCCTTACAGAGCTGCTTC	1139	
QY	1157	CGCCAGATTCCGTGAGCC	CCCTTTATGGGCTTCA	CTTCAACATCAGCCTGAAACAAGA	1216
Db	1140	CGCCAGATTCCGTGAGCC	CTTATGGGCTTCA	CTTCAACATCAGCCTGAAACAAGA	1199
QY	1217	GGAATGAGATTCCAGCCCAAT	CAGAGAGACAGAGAGATCTC	ACGCTGAGCATCATTTG	1276
Db	1200	GGAATGAGATTCCAGCCCAAT	CAGAGAGACAGAGAGATCTC	ACGCTGAGCATCATTTG	1259
QY	1277	CGCCTTCCTAAGGCGT	GAGTCCAGATCA	CATCTTCCAGGGCAACCTCAAGAGCAAA	1336
Db	1260	CGCCTTCCTAAGGCGT	GAGTCCAGATCA	CATCTTCCAGGGCAACCTCAAGAGCAAA	1319
QY	1337	ACTACTGTGGCCCAAGAGGAA	TCCCTTCTCA	CGAGGGGCTGCGCCAAAACAAGAGC	1396
Db	1320	ACTACTGTGGCCCAAGAGGAA	TCCCTTCTCA	CGAGGGGCTGCGCCAAAACAAGAGC	1379
QY	1397	AGCCAAACAGAACGTTA	GGGGCCAGAGAACAAACAGGCTCGAAAGTTAAGCTTGTGA	1456	
Db	1380	AGCCAAACAGAACGTTA	GGGGCCAGAGAACAAACAGGCTCGAAAGTTAAGCTTGTGA	1439	
QY	1457	CGCCTTCAAGTCTGGCCCACT	GTATCAGAGGCCAGGCTA	CTAAGTGGCCCCACAGAGCGC	1516
Db	1440	CGCCTTCAAGTCTGGCCCACT	GTATCAGAGGCCAGGCTA	CTAAGTGGCCCCACAGAGCGC	1499
QY	1517	CCTGAGCCCCACTCCAT	GTGTTCTTCCCCCTTAA	CAATCAACGCGCGTCAAAAGTTTCACAG	1576
Db	1500	CCTGAGCCCCACTCCAT	GTGTTCTTCCCCCTTAA	CAATCAACGCGCGTCAAAAGTTTCACAG	1559
QY	1577	TGTCAACAGCATAGAC	CAAAAGCAAAAGCTTAA	AACTGTGAAGTCTTGGGGCCAAAGA	1636
Db	1620	AAAGTTGAATGTCTCA	GAGAGCGATGGGGCTTGA	GAGCAACCGAAGTATCTCA	1696
QY	1697	AGTGAAGAGAAAACTGT	GAGTTTAACTGACGATATG	CCAGAGATTCGCCGAAATCA	1756
Db	1680	AGTGAAGAGAAAACTGT	GAGTTTAACTGACGATATG	CCAGAGATTCGCCGAAATCA	1739
QY	1757	CCTTAAAGAACTTTGGAA	CAATACCAACCAATACACTA	CACTCAACTCAAGATTCACAT	1816
Db	1740	CCTTAAAGAACTTTGGAA	CAATACCAACCAATACACTA	CACTCAACTCAAGATTCACAT	1799
QY	1817	GGAATCCTTATTTGGG	CGTTGAAAAACAGGGATGA	ACACATTCCTAACCTGCTTCTTATG	1876
Db	1800	GGAATCCTTATTTGGG	CGTTGAAAAACAGGGATGA	ACACATTCCTAACCTGCTTCTTATG	1859
QY	1877	GGGATGCTTGGCCAGC	AGTCTTCACTGTGTGTACA	CCAGACAGACACTGATTCAGTGC	1936
Db	1860	GGGATGCTTGGCCAGC	AGTCTTCACTGTGTGTACA	CCAGACAGACACTGATTCAGTGC	1919
QY	1937	ACAGCATTCAGCTGTCA	CACTGAAGAAAGCTGTCTT	CAACAGCCTGGAATCAATGGTT	1996
Db	1920	ACAGCATTCAGCTGTCA	CACTGAAGAAAGCTGTCTT	CAACAGCCTGGAATCAATGGTT	1979

QY	1997	AGCTTAATAGATAAAATCCAGACTTACTTACGCTTTATATGCTTTATTATCAATAAAC	2056
Db	1980	AGCTTAATAGATAAAATCCAGACTTACTTACGCTTTATATGCTTTATTATCAATAAAC	2039
QY	2057	TGTGAACCTGACGTAACCAATGGAAACATTATACAGACTCTGGAATTCAGACTCGG	2116
Db	2040	TGTGAACCTGACGTAACCAATGGAAACATTATACAGACTCTGGAATTCAGACTCGG	2099
QY	2117	AACCTTAGTCTATCTGATCCAGACAGACACACTTAGTATCTGCCAACTAATG	2176
Db	2100	AACCTTAGTCTATCTGATCCAGACAGACACACTTAGTATCTGCCAACTAATG	2155
QY	2177	AGTTAATPAA	2187
Db	2160	AGTTAATPAA	2170
RESULT 5			
LOCUS	AY357925	2187 bp	mRNA
DEFINITION	Macaca faecicularis beestrophin (VMD2) mRNA, complete cds.	linear	PRI 26-AUG-2003
ACCESSION	AY357925		
VERSION	AY357925.1	GI:34013782	
KEYWORDS			
SOURCE	Macaca faecicularis (crab-eating macaque)		
ORGANISM	Macaca faecicularis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae; Cercopitheciinae; Macaca.		
AUTHORS	1 (bases 1 to 2187) Okamoto,H., Umeda,S., Suzuki,M.T., Yoshihawa,Y., Tanaka,Y. and Iwata,Y.		
TITLE	Molecular Cloning of VMD2 Gene from Cynomolgus Monkey (Macaca faecicularis)		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2187)		
AUTHORS	Okamoto,H., Umeda,S., Suzuki,M.T., Yoshihawa,Y., Tanaka,Y. and Iwata,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-AUG-2003) National Tokyo Medical Center, National Institute of Sensory Organs, 2-5-1, Higashiagaoka, Meguro-ku, Tokyo 152-0021, Japan		
FEATURES			
source	Location/Qualifiers 1..2187 /organism="Macaca faecicularis" /mol_type="mRNA" /db_xref="taxon:9541" 1..2187 /gene="VMD2" 67..1824 /gene="VMD2" /codon_start=1 /product="beestrophin" /protein_id="AA056049.1" /db_xref="GI:34013783" /translation="MTITTSQVANARIGSTSRLLCMKRGITIKLLIGEPFIPILCYYY IIRIRLALTEEQQLMEKRLTYLSDSYLIPLISFVLGFTVTLVTRMMQYNLPLP IDRLMSLVGVEQDEQGRLLRLRLRYANIGVLLIRSVTVYKRFPSAQHLYOAH GPMPLAEHKQLEKLSLPHNMFWPMWPMANSMKAMGGRJRDPILOSILNENLTLE TQGHLLAYDMISIPLYVTQVTVAAVVSFFPLCLVGRQFLIPAKAYBGEHLDVVPV TFLEFFEVGMLKYAEOLINRGEDDDPEFNMTVDRLQVSLAIDEMODLPRMER DMYNEPRPHRYTAAKQPPRASMGSTFNISLNKEMEQPNQOEDKEDATGIIIGER FLGQSDHHPPGANSRTKLLMPKRESLLHSGLEPNKAVIQNVRGLEDNAAMLKAK DAFKSAPLYQRPGRYASAPQTPLSPMPFPPEPSVLSGLHSTGVIDTKSLKLVSSG AKNSSELISGDMERPEVSHVARKTVEFNLTDMPEIPIINHLEKEPLEISTYIMATW LKDHVPYMALENRDEAHS"		
gene			
CDS			
ORIGIN			
Query Match	90.7%;	Score 2022.6;	DB 9; Length 2187;
Best Local Similarity	95.7%;	Pred. No. 0;	
Matches 2092;	Conservative	0; Mismatches 89;	Indels 6; Gaps 1;

QY	39	GGATCATGGAAACCAACCTGGAAACCCACCTGACCCAAAGCCCACTGCTGGACCCCACTGC	98
Db	1	GGATCACCGGATCCACTGGAAACCCCACTGACCCCAAGCCCACTGCTGGACCCAAAGTGC	60
QY	99	CTGGCCATGACCATCATCTTACACAAAGCAATGGCTAAATGGCCGTTAGCTCTCTTCTCC	158
Db	61	CTGGCCATGACCATCTTACACAGAGCCAAATGGCTAAATGCCGTTAGCTCTTCTCTCC	120
QY	159	CGCCTGCTGTGTGTGGCGGGGAGCATCTACAAGCTGTATATATGCGAGTTCTTAAATC	218
Db	121	CGCCTGCTGTGTGTGGCGGGGAGCATCTACAAGCTGTATATGCGAGTTCTTCAATC	180
QY	219	TTCTGCTCTGTCTACTATCATATCCGCTTTATTTATAGCTGGCCCTTACGGAAGAACAA	278
Db	181	TTCTGCTCTGTCTACTATCATATCCGCTTTATTTATAGCTGGCCCTTACAGAAAGAACAA	240
QY	279	CAGCTGATGTTTGAAACATGACTCTGATTTGGGACGACTATCATCCAGCTATCCCATTT	338
Db	241	CAGCTGATGTTTGAAAGACTGACTCTGATTTGGGACGACTATCATCCAGCTATCCCATTT	300
QY	339	TCCTTCTGTGTGGGCTTCTACGTGACGCTGTGTGTGACCCGCTGGTGGACACAGTACAG	398
Db	301	TCCTTCTGTGTGGGCTTCTACGTGACGCTGTGTGTGACCCGCTGGTGGACACAGTACAG	360
QY	399	AACTGCTGGTGGCCGACACCGCTCATGAGCTGTGTGGGCTTGTGTGAAGCGAAGAC	458
Db	361	AACTGCTGGTGGCCGACACCGCTCATGAGCTGTGTGTGGGCTTGTGTGAAGCGAAGAC	420
QY	459	GAGCAAGCGGGGCTGTGCGGGGCGACGCTATTCGGCTACCGCCAACTGGGGAACGCTGTC	518
Db	421	GAGCAAGCGGGGCTGTGCGGGGCGACGCTATTCGGCTACCGCCAACTGGGGAACGCTGTC	480
QY	519	ATCTGTGCGACGCTCAGACACCGCAGTCTTACAAGGCGCTTCCCGAGCGCCAGACACTGTGTG	578
Db	481	ATCTGTGCGACGCTCAGACACCGCGGCTTACAAGGCGCTTCCCGAGCGCCAGACACTGTGTG	540
QY	579	CAAGCAGGCTTTATGACTCCGGCGAGAACACAGAGTTGGAGAACTGAGCTTACACAC	638
Db	541	CAAGCAGGCTTTATGAGACCCCGCGAGAACACAGAGTTGGAGAACTGAGCTTACACAC	600
QY	639	AAACAATGTTCTGGGGTGCCCTGGGGTGTGGTTGCCAACTGTAAATGAAGGCGTGGCTTGG	698
Db	601	AAACAATGTTCTGGGGTGCCCTGGGGTGTGGTTGCCAACTGTAAATGAAGGCGTGGCTTGG	660
QY	699	GGTGGAATCCGGGACCCCTATCTGTCTCCAGAGCTGTGCTGAACGAGATGAACACTTGTGCT	758
Db	661	GGTGGAATCCGGGACCCCTATCTGTCTCCAGAGCTGTGCTGAACGAGATGAACACTTGTGCT	720
QY	759	ACTGAGTGTGACACTGTGTATGCTTACGACTGTGATTAAGTATCCCATGTGTATACAG	818
Db	721	ACTGAGTGTGACACTGTGTATGCTTACGACTGTGATTAAGTATCCCATGTGTATACAG	780
QY	819	GTGTGTACTGTGGGGGTGTACAGCTTCTTCTGTACTTGTCTAATTTGGGCGGCAATTTCTG	878
Db	781	GTGTGTACTGTGGGGGTGTACAGCTTCTTCTGTACTTGTCTAATTTGGGCGGCAATTTCTG	840
QY	879	AAACCAGGCAAGGCGCTACCTCGGCGCATGAGCTGGAACCTCGTTGTGGCCGCTTTCAAGTTT	938
Db	841	AAACCAGGCAAGGCGCTACCTCGGCGCATGAGCTGGAACCTCGTTGTGGCCGCTTTCAAGTTT	900
QY	939	CTGCAAGTTCTTCTTCTATGTGTGGCTGAGGTGGGAGAGCAAGCTCATCAACCCCTTT	998
Db	901	CTGCAAGTTCTTCTTCTATGTGTGGCTGAGGTGGGAGAGCAAGCTCATCAACCCGTTT	960
QY	999	GGAGAAGATGATGATTTTGAACAACAATGTGTGCAACAGAAATTTGCAAGTGTCC	1058
Db	961	GGAGAAGACGATGATTTTGAACAACAATGTGTGCAACAGAAATTTGCAAGTGTCC	1020
QY	1059	CTGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1118
Db	1021	CTGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1080
QY	1119	AATAGGCCGAGCCAGCCGCCCTTACAGACTGCTTCCGCCAGTTCCGTGAGCCCTCC	1178

Db	1081	AATATAGCTTAGGCCACACCCCCCTTACAGCTGCTTCGGCCAGTTCCGTGACCTCC	1140
Qy	1179	TTTATGGGCTCCACCTTCAACATCAAGCTGGAACAAAGAGATGGAAGTTCCAGCCCAAT	1238
Db	1141	TTTATGGGCTCCACCTTCAACATCAAGCTGGAACAAAGAGATGGAAGTTTCAAGCCCAAT	1200
Qy	1239	CAGAGAGACGAGGAGATGCTCAAGCTGACATTTGGCCGCTTCTTAGGCTGTCAATCC	1298
Db	1201	CAGAGAGACAAAGAGACGCTCAACTGGCATCATTTGGCCGCTTCTTAGGCTGTCAATCC	1260
Qy	1299	CATGATCAACATCTTCCAGGGCAACTCAAGAGCAAACTACTGTGCCCCAAGAGGGA	1358
Db	1261	CATGATCAACATCTCTCCGGGGCAACTCAAGAGCAAACTACTGTGCCCCAAGAGGGA	1320
Qy	1359	TCCCTTCTCAGAGGGGCTCGCCAAAAACCAAGGCAAGCCAAACAGACGTTAGGGGC	1418
Db	1321	TCCCTTCTCAGAGGGGCTCGCCAAAAACCAAGGCAAGCCAAACAGACGTTAGGGGC	1380
Qy	1419	CAGAAAGACAAAGGCTGGAAGCTTAAAGCTGTGGAAGCCCTTCAAGTCTGAGCCACTG	1478
Db	1381	CTGAAGACAAACAGGCTGGAAGCTTAAAGCTGTGGAAGCTTCAAGTCTGAGCCACTG	1440
Qy	1479	TATCAAGGCGAGGCTACTACAGTGGCCCCACAGAGGCCCTTCAAGCCCACTCCATGTC	1538
Db	1441	TATCAAGGCGAGGCTACTACAGTGGCCCCACAGAGGCCCTTCAAGCCCACTCCATGTC	1500
Qy	1539	TTCCCTCTGAACCAATCAGCGCGCTCAAGGCTTCAAGTGTCAAGAGCATATGACCCAA	1598
Db	1501	TTCCCTCTGAACCAATCAGTGTGTCAAGGCTTCAAGTGTCAAGAGCATATGACCCAA	1560
Qy	1599	GACAAAAGCTTAAAGCTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAG	1658
Db	1561	GACAAAAGCTTAAAGCTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAG	1620
Qy	1659	AACGATGGGGCTTATATGGAGACCCCAAACTATCTCAAGTGAAGAGAAAATGTGGAG	1718
Db	1621	AACGATGGGGCTTATATGGAGACCCCAAACTATCTCAAGTGAAGAGAAAATGTGGAG	1680
Qy	1719	TTTAACTGAAGATATGCAAGATCCCGAAATCACTCAAGAACTTTGGAAACAA	1778
Db	1681	TTTAACTGAAGATATGCAAGATCCCGAAATCACTCAAGAACTTTGGAAACAA	1740
Qy	1779	TCACCAACCAACATACATCACTACACTCAAGATCACTGATCTCTATTGGGCTTGGAA	1838
Db	1741	TCACCAACCAACATACATCACTACACTCAAGATCACTGATCTCTATTGGGCTTGGAA	1800
Qy	1839	AAACAGGATGGAAGCAATCTCTAACTGCTTCTTAATGGGGATGCTTCGCAAGCAGATC	1898
Db	1801	AAACAGGATGGAAGCAATCTCTAACTGCTTCTTAATGGGGATGCTTCGCAAGCAGATC	1860
Qy	1899	CTCAACTGTGTGTACACGACGAGACACTGATTCACAGCCATACAGCTGTCCACAC	1958
Db	1861	CTCTCCGTTGTGTACACGACGAGACACCGATTCAGCTACAGCCATACAGCTGTCCACAC	1920
Qy	1959	TGAAGAACTGTCTCTTCAACAGCCTGAAATCAAAATGGTGGCTTAATAGATAAA-----A	2012
Db	1921	TGAAGAACTGTCTCTTCAACAGCCTTCAATGGTGGCTTAATAGATAAAATCAGCA	1980
Qy	2013	ATCCCAAGCTACTTCAAGCTTTTAATGCTTTTATTCATAAAAATGTGAAAGCTAGACTG	2072
Db	1981	GTCACAGACTACTTCAAGCTTTTAATGCTTTTATTCATAAAAATTAAGAACTAGACTG	2040
Qy	2073	AACCATGGAAACATTTAACTCAGACTCTGGAACTTCAGAGTCGGGAACCTTAGTCTATTC	2132
Db	2041	AACCATGGAAACATTTAGCTCAGACCTGGAACTTGAAGTTCAGGAACCTTTATCTCTGTC	2100
Qy	2133	TGAATCCAAAGACGCCACACTTAGTATATCTGGCCAAACTAATGAGTTTAAATAATCAA	2192
Db	2101	TGAATCCAAAGACGCCATACCTCAGATATCTGGCCAAACTAAGGCGTTTAAATAATCAA	2160
Qy	2193	ATTCTGGTTAAAAAATAAAAAATAAAAA 2219	

DB	2161	ATACCTGTTTAAAAAAAAAAAAA	2187
RESULT 6			
LOCUS	BD136721	2429 bp	DNA linear PAT 18-SEP-2002
DEFINITION	Best's macular dystrophy gene.		
ACCESSION	BD136721		
VERSION	BD136721.1 GI:23231666		
KEYWORDS	JP 2002504559-A/3.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 2429)		
JOURNAL	Petrushin,K., Caskey,T.C., Metzker,M. and Wadelius,C. Best's macular dystrophy gene Patent : JP 2002504559-A 3 12-FEB-2002; MERCK & CO INC,CLAES WADELIIUS		
COMMENT	OS Homo sapiens (human) PN JP 2002504559-A/3 PD 12-FEB-2002 PP 22-FEB-1999 JP 2000533447 PR 25-FEB-1998 US 60/075941,18-DEC-1998 US 60/112926 PI CONSTANTIN PETRUSHIN,THOMAS C CASKEY,MICHAEL METZKER,CLAES PI WADELIIUS PC C07K16/18,C07K14/47,C12N5/10,C12N15/09,C12P19/34,C12Q1/68// PC CI2P11/08, PC C12N5/00,C12N15/00 CC Best's macular dystrophy gene FH Key Location/Qualifiers FT source 1..2429 Location/Qualifiers 1..2429 /organism='Homo sapiens' /mol_type='genomic DNA' /db_xref='taxon:9606'		
FEATURES			
source			
ORIGIN			
Query Match	90.3%:	Score 2013;	DB 6; Length 2429;
Best Local Similarity	91.6%:	Pred. No.0;	
Matches 2226;	Conservative 0;	Mismatches 0;	Indels 203; Gaps 1;
OY	1	CAGGAGTCGCCAGCAGCTAGTGCGGCAGACCCTTCTGTGGATCATGGAACCACTGGAA	60
Db	1	CAGGAGTCCACAGCAGCTAGTGCGGCAGACTTCTGTGGATCATGGAACCACTGGAA	60
OY	61	CCCCACCTGACCCAGGCCACCTGCTGACGCCACCTGCTGCTGATGACCATCATTA	120
Db	61	CCCCACCTGACCCAGGCCACCTGCTGACGCCACCTGCTGCTGATGACCATCATTA	120
OY	121	CAAACCAAGTGAGTAATGACCGCTTAGGCTCTTCTCCGCGCTGCTGCTGCTGCTGCGGG	180
Db	121	CAAACCAAGTGAGTAATGACCGCTTAGGCTCTTCTCCGCGCTGCTGCTGCTGCTGCGGG	180
OY	181	GACGATCTCAAGCTGCTATATGCGAGGTTCTTAATCTTCTGCTGCTGCTACTACATCA	240
Db	181	GACGATCTCAAGCTGCTATATGCGAGGTTCTTAATCTTCTGCTGCTGCTACTACATCA	240
OY	241	TCCGCTTTATTATAGGCTGCGCCCTCAGAGGAACAAGCTGATGTTTGGAAATCGA	300
Db	241	TCCGCTTTATTATAGGCTGCGCCCTCAGAGGAACAAGCTGATGTTTGGAAATCGA	300
OY	301	CTCTGATTTGCGACAGCTACATCACTCACTCATCCCATTTCTTCTGCTGAGGCTTCTACG	360
Db	301	CTCTGATTTGCGACAGCTACATCACTCACTCATCCCATTTCTTCTGCTGAGGCTTCTACG	360
OY	361	TGAGCGTGGCTGCGAACCCGCTGTGTGAACCAATGAGAATCTGCGTGGCCGACGCC	420
Db	361	TGAGCGTGGCTGCGAACCCGCTGTGTGAACCAATGAGAATCTGCGTGGCCGACGCC	420
OY	421	TCAATGAGCTGTGTGTGGGCTTCTGTGAAGCAAGCAAGAGCAAGGCGAGCTGCTGAGGC	480

D	b	421	TATGAGCCTTGCTGTGCGCTTCTGTGAAGGCAAGACGAGCAAGGCCAGCTCTGGC	480
O	y	481	GCACGCTCATCCGCTACGCGCACTTGAGCACTGGCTCATCTCGCACGTCGAGACCG	540
D	b	481	GCACGCTCATCCGCTACGCGCACTTGAGCACTGGCTCATCTCGCACGTCGAGACCG	540
O	y	541	CAGTCTACAGAGGCTTCCCGAGGCCGACGACTGGTGCAAGAGGCTTATGATCCGG	600
D	b	541	CAGTCTACAGAGGCTTCCCGAGGCCGAGACTGGTGCAAGAGGCTTATGATCCGG	600
O	y	601	CAGAACACAAAGGACTTGGAGAACTGAGCTTACACACAATGTTCTGGGTGCCCTGGG	660
D	b	601	CAGAACACAAAGGACTTGGAGAACTGAGCTTACACACAATGTTCTGGGTGCCCTGGG	660
O	y	661	TGTGTTTTCGCAACTGTCAATGAAAGGCGCTTGGAGGTGCAGATCCGAGACCTATCC	720
D	b	661	TGTGTTTTCGCAACTGTCAATGAAAGGCGCTTGGAGGTGCAGATCCGAGACCTATCC	720
O	y	721	TGCTCCAGAGCCTGCTGAAAGAGATGAACCTTGGCTACAGTGCTGACACTGTATG	780
D	b	721	TGCTCCAGAGCCTGCTGAAAGAGATGAACCTTGGCTACAGTGCTGACACTGTATG	780
O	y	781	CCTAGGACTGGAATTAGTATCCCATGTGTATACAGAGTGTGACTGTGGCGGTGACA	840
D	b	781	CCTAGGACTGGAATTAGTATCCCATGTGTATACAGAGTGTGACTGTGGCGGTGACA	840
O	y	841	GCTTCTTCTGCACTTGTAGTTGGGCGGCACTTTCGAAACCGACGCAAGGCTTACCTG	900
D	b	841	GCTTCTTCTGCACTTGTAGTTGGGCGGCACTTTCGAAACCGACGCAAGGCTTACCTG	900
O	y	901	GCATGAGCTGGAACCTCGTGTGTGCGCCGCTTTCACGTTCTGCAATTCCTTATGTTG	960
D	b	901	GCATGAGCTGGAACCTCGTGTGTGCGCCGCTTTCACGTTCTGCAATTCCTTATGTTG	960
O	y	961	GCTGGCTGA-----	970
D	b	961	GCTGGCTGAAGGTGGGCTCTCCAGGGCCCTGTGGGCTGAGGCAATGGCCAGAGGGATC	1020
O	y	971	-----	970
D	b	1021	ATGGCCAGCACTGCTTGAAGACGAGATGCAAGTTCAGAAAGAAAGTCTCACGGGTAG	1080
O	y	971	-----	970
D	b	1081	AAAGCAGCCAGGCGTGTGGCGCACACTGTATCCAGCTACTCGGTAAGCTGAGAGCAG	1140
O	y	971	-----GTTGGCAGAGCAGCTCATCAACCCCTT	997
D	b	1141	GAGAAATCGCTTGAACCCGGAAGCGAGGTGTGTGGCAGAGCAACTCATCAACCCCTT	1200
O	y	998	TGGAGAGGATGATGATTTTGAAGCAAACTGGATTGTGCACAGAAATTTTGAAGGTTC	1050
D	b	1201	TGGAGAGGATGATGATTTTGAAGCAAACTGGATTGTGCACAGAAATTTTGAAGGTTC	1260
O	y	1058	CCTGTGTGCTGTGATGATGATGACCAAGGACCTGCTCGGATGGAGCAGGACATGTACTG	1110
D	b	1261	CCTGTGTGCTGTGATGATGATGACCAAGGACCTGCTCGGATGGAGCAGGACATGTACTG	1320
O	y	1118	GAAATAGCCGAGCCACAGCCCCCTTACACAGCTGCTTCGCCAGTTCGTGAGACCTC	1170
D	b	1321	GAAATAGCCGAGCCACAGCCCCCTTACACAGCTGCTTCGCCAGTTCGTGAGACCTC	1380
O	y	1178	CTTTATGGGCTCACCTTCAACATCGAGCTGAAACAAAGAGAGATGAGATTCAGGCCAA	1230
D	b	1381	CTTTATGGGCTCACCTTCAACATCGAGCTGAAACAAAGAGAGATGAGATTCAGGCCAA	1440
O	y	1238	TCAGAGAGGACGAGAGGATGCTCACCGTGGCATATTTGACCGCTTCTTAGGCTCGAGTC	1290
D	b	1441	TCAGAGAGGACGAGAGGATGCTCACCGTGGCATATTTGACCGCTTCTTAGGCTCGAGTC	1500
O	y	1298	CCATGATCACATCTCTCCAGAGCAAACTCAAGGACCAACTATCTGTGGCCCAAGAGGA	1350

Db	1501	CCATGATACCATCTCTCCACAGGCGAACTCAAGAACCAAACTACTGTGCGCCCAAGAGGGA	156
QY	1358	ATCCCTCTCCACGAGGGCGCTGCCCAAAACACAGAAGGACGCAACGAAAGCTTAAAGGG	1417
Db	1561	ATCCCTCTCCACGAGGGCGCTGCCCAAAACCAACAGGACGCAACAGAAAGCTTAAAGGG	1620
QY	1418	CCAGGAAGACAACAAAGGCTGTGAAAGCTTAAAGCTGTGGAAGCCCTTCAAGTCTGAGCCCACT	1477
Db	1621	CCAGGAAGACAACAAAGGCTGTGAAAGCTTAAAGCTGTGGAAGCCCTTCAAGTCTGAGCCCACT	1680
QY	1478	GTATCAGAGGCGACGACTACTACAGTGCAGCCCAAGACGCGCCCTCAGCGCCCACTCCCATGTT	1537
Db	1661	GTATCAGAGGCGACGACTACTACAGTGCAGCCCAAGACGCGCCCTCAGCGCCCACTCCCATGTT	1744
QY	1538	CTTCCCTCCTAGAACCATCAGCGCGCGTCAAAGCTTCAAGTGTCAAGGCATAGACCA	1597
Db	1741	CTTCCCTCCTAGAACCATCAGCGCGCGTCAAAGCTTCAAGTGTCAAGGCATAGACCA	1800
QY	1558	AGACAAAAGCTTAAAGACTGTGAAGTCTTGAGGGCCCAAGAAAAGTTTGAATTGCTCTCAGA	1657
Db	1801	AGACAAAAGCTTAAAGACTGTGAAGTCTTGAGGGCCCAAGAAAAGTTTGAATTGCTCTCAGA	1860
QY	1658	GAGCGATGGGGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGGAAGAAAGCTGTGGA	1717
Db	1861	GAGCGATGGGGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGGAAGAAAGCTGTGGA	1920
QY	1718	GTTTAACTCAGCGATATGCGCAGAGATCCCGGAAATATCACTCAAGAACCTTTGGAACA	1777
Db	1921	GTTTAACTCAGCGATATGCGCAGAGATCCCGGAAATATCACTCAAGAACCTTTGGAACA	1980
QY	1778	ATCACCAACCAACATACACTACATCACTCAAGATACATGATCTCTTATTTGGGCTTTGGA	1837
Db	1981	ATCACCAACCAACATACACTACATCACTCAAGATACATGATCTCTTATTTGGGCTTTGGA	2040
QY	1838	AAACAGGATGAACACACTTCTTAACTCGCTCTCTTAAATGGGGAGTCTTCCGACGCAAGST	1897
Db	2041	AAACAGGATGAACACACTTCTTAACTCGCTCTCTTAAATGGGGAGTCTTCCGACGCAAGST	2100
QY	1898	CCTCACCTGTGTGTAACACAGCAGAGACACTGATCCAGTCAAGGCATATCAGCTGTCCACA	1957
Db	2101	CCTCACCTGTGTGTAACACAGCAGAGACACTGATCCAGTCAAGGCATATCAGCTGTCCACA	2166
QY	1958	CTGAAGAACGTGTCTTACACACAGGCTGTGAATCAAAATGTTAGCTTAAATAGATAAAATATCC	2017
Db	2161	CTGAAGAACGTGTCTTACACACAGGCTGTGAATCAAAATGTTAGCTTAAATAGATAAAATATCC	2220
QY	2018	AGACTACTCTTACGCTTTAATGCTTTTATATATATAAAACTGTGAAGCTAGACTGAACCA	2077
Db	2221	AGACTACTCTTACGCTTTAATGCTTTTATATATATAAAACTGTGAAGCTAGACTGAACCA	2280
QY	2078	TTGGAAACATTTAATCTCAGACTCTGGAATTCAGAGTCGGGAAACCTTATGTTCTATCTGAAT	2137
Db	2281	TTGGAAACATTTAATCTCAGACTCTGGAATTCAGAGTCGGGAAACCTTATGTTCTATCTGAAT	2340
QY	2138	CCAAAGACGACACCTTATGATATCTGCCCAAACTAAATGAGTTTAAATATACAAATATCT	2197
Db	2341	CCAAAGACGACACCTTATGATATCTGCCCAAACTAAATGAGTTTAAATATACAAATATCT	2400
QY	2198	CGTTAAAAAAAAAAAAAAAAAAAAA 2226	
Db	2401	CGTTAAAAAAAAAAAAAAAAAAAAA 2429	

RESULT 7	AF057170	2420 bp	mRNA	linear	PRI 17-OCT-1998
LOCUS	AF057170				
DEFINITION	AF057170 Homo sapiens beetrophin (VMD2) mRNA, alternatively spliced product,				
ACCESSION	AF057170				
VERSION	AF057170.1	GI:3335160			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euleleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 2420)

Petrushin, K., Koisti, M.J., Bakall, B., Li, M., Xie, G., Marknell, T., Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M., Bergen, A.A., McArthur, D., Figueras, D., Austin, C.P., Metzker, M.L., Caskey, C.T. and Madeline, C. Identification of the gene responsible for Best macular dystrophy Nat. Genet. 19 (3), 241-247 (1998)

98324772

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

2 (bases 1 to 2420)
Petrushin, K.
Direct Submission
Submitted (03-APR-1998) Human Genetics, Merck Research Laboratories, West Point, PA 19486, USA
Location/Qualifiers
1..2420

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..2420
/gene="VMD2"
/gene="VMD2"
/note="Best's macular dystrophy gene; BMD, alternatively spliced isoform; contains exon 7"
/product="bestrophin"
/protein_id="AAC64344.1"
/db_xref="GI:3335161"
/translation="MTITVTSQVANARIQSFSRLILCWRSIYKLLYGEPLIPILCYV
IIRFYRLALTEBOQLMEKLLYCDSTIOLIPISFVLGFTVLTVMRMNOVNLPM
PRMLSVLSFVEKQBOGRLRLRLIYANLGNVLILRSSTAVYRFPRAQNVQA
GMPRAHEKOLEKSLPHNMFPVPMVPMANISKAYLGRJRDPILOSILNENVTVA
TQCGHLAYADMSIPLVYTVYVAVISFPLCLVGRQFLNPAAYPHELDLIVPVF
TFLQFFVYGMILKGLSRALLGWRHGGHGOQLLETMOCEKRVSVESQAMWPT
PVIPATRAEAGESELEPGRRLIMQSSSTPLEMMMLIRPTGSTGICRCPCVLMWR
CTRTCLGMSRCTGISPSHSPTQLPPSSVPEPLMAPPSRSA"

ORIGIN

Query Match 89.8%; Score 2002.4; DB 9; Length 2420;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 2216; Conservative 0; Mismatches 1; Indels 203; Gaps 1;

QY 1 CAGGAGTCCACAGCCTTGTGCGAGACCTTCTGTGGATCATCGGACCCACTTGA 60
Db 1 CAGGAGTCCACAGCCTTGTGCGAGACCTTCTGTGGATCATCGGACCCACTTGA 60
QY 61 CCCACCTGACCCCAAGCCCACTGTGTGAGCCCACTGTGGCATGACATCACTTACA 120
Db 61 CCCACCTGACCCCAAGCCCACTGTGTGAGCCCACTGTGGCATGACATCACTTACA 120
QY 121 CAAGCAAGTGGCTAATGCCGCTTAAAGGCTCTCTCCCGCTGCTGTGTGGGGGG 180
Db 121 CAAGCAAGTGGCTAATGCCGCTTAAAGGCTCTCTCCCGCTGCTGTGTGGGGGG 180
QY 181 GCAGCATCTACAACTGTATATGCGAGTCTTAACTCTCTGCTGTGCTACTACATCA 240
Db 181 GCAGCATCTACAACTGTATATGCGAGTCTTAACTCTCTGCTGTGCTACTACATCA 240
QY 241 TCGCCTTATTTAATAGCTGGCCCTCAAGCAAGCAAGCTGATGTTGAAGAACTGA 300
Db 241 TCGCCTTATTTAATAGCTGGCCCTCAAGCAAGCAAGCTGATGTTGAAGAACTGA 300
QY 301 CTCTGATTTGAGAGCTAGATCAAGCTATCCCATTTCTTCTGAGTGGGCTTTACG 360
Db 301 CTCTGATTTGAGAGCTAGATCAAGCTATCCCATTTCTTCTGAGTGGGCTTTACG 360
QY 361 TGAAGCTGTGTGAGACCCGCTGTGGAACAGATGAGAACTGTCCGTGGCCGAGCGCC 420
Db 361 TGAAGCTGTGTGAGACCCGCTGTGGAACAGATGAGAACTGTCCGTGGCCGAGCGCC 420

Db 361 TGAAGCTGTGTGAGACCCGCTGTGGAACAGATGAGAACTGTCCGTGGCCGAGCGCC 420
QY 421 TCATGAGCTGTGTGAGAGCTTTCGTCGAAGGCAAGCAAGCCGCTGTGGGGC 480
Db 421 TCATGAGCTGTGTGAGAGCTTTCGTCGAAGGCAAGCAAGCCGCTGTGGGGC 480
QY 481 GCAGCTCATTCGCTACGCGCAACCTGGCAAGCTGTATCTGCGAGCGTACGACCG 540
Db 481 GCAGCTCATTCGCTACGCGCAACCTGGCAAGCTGTATCTGCGAGCGTACGACCG 540
QY 541 CAGTCTACAGCGCTTCCCGAGCGCCAGCACTGTGTGCAAGGCTTTATGACTCCG 600
Db 541 CAGTCTACAGCGCTTCCCGAGCGCCAGCACTGTGTGCAAGGCTTTATGACTCCG 600
QY 601 CAGAACACAGCACTGTGAGAACTGAGCTTACCAACAATGTTGTGGGTGGCCTGAG 660
Db 601 CAGAACACAGCACTGTGAGAACTGAGCTTACCAACAATGTTGTGGGTGGCCTGAG 660
QY 661 TGTGTTTTCGCAACCTGTCAATGAAAGCGCTTGGAGGTCAATCCGGACCTATCC 720
Db 661 TGTGTTTTCGCAACCTGTCAATGAAAGCGCTTGGAGGTCAATCCGGACCTATCC 720
QY 721 TGTCTCAGAGCTGTCTGAAAGAGTGAACCTTGGCTACTCACTGTGCACTGTATG 780
Db 721 TGTCTCAGAGCTGTCTGAAAGAGTGAACCTTGGCTACTCACTGTGCACTGTATG 780
QY 781 CTTACGACTGATTAATATCCACTGTGTATACAGAGTGTGACTGTGGCGGTATAG 840
Db 781 CTTACGACTGATTAATATCCACTGTGTATACAGAGTGTGACTGTGGCGGTATAG 840
QY 841 GCTTCTTCTGACTGTCTGATGATGTTGGGCGAGTTCGAAACCAAGCAAGCCTTACCTG 900
Db 841 GCTTCTTCTGACTGTCTGATGATGTTGGGCGAGTTCGAAACCAAGCAAGCCTTACCTG 900
QY 901 GCATGAGCTGAGACTGTGTGTCGCCGTCTTCACTGTGCAAGTCTTCTATATGTTG 960
Db 901 GCATGAGCTGAGACTGTGTGTCGCCGTCTTCACTGTGCAAGTCTTCTATATGTTG 960
QY 961 GCTGGCTGAA----- 970
Db 961 GCTGGCTGAAAGTGGGCTCTTCCAGGGCCCTGTGGGCTGAGGCAATGCGAGGGGCTC 1020
QY 971 ----- 970
Db 971 ATGGCAGCAGCTGCTTGAAGAGAGATGCAATGTCAGGAAGAGAGTCTTCAAGGCTT 1080
QY 971 ----- 970
Db 971 AAAGCAGCAGGCTGTGTGCGCAGACCTGTAATCCAGCTACTCGGAGGCTGAGGAG 1140
QY 971 -----GGTGGCAGAGCGCTCATCAACCTCTT 997
Db 971 GAGATCGCTTAAACCGGAGGCGAGAGGTGTGTGCGAGCGCTCATCAACCTCTT 1200
QY 998 TGAAGAGATGATGATGATTTTGAAGCACTGATTTGCAAGAGATTTGCAAGTTC 1057
Db 998 TGAAGAGATGATGATGATTTTGAAGCACTGATTTGCAAGAGATTTGCAAGTTC 1057
QY 1201 TGAAGAGATGATGATGATTTTGAAGCACTGATTTGCAAGAGATTTGCAAGTTC 1260
Db 1201 TGAAGAGATGATGATGATTTTGAAGCACTGATTTGCAAGAGATTTGCAAGTTC 1260
QY 1058 CCTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1117
Db 1058 CCTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1117
QY 1261 CCTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db 1261 GAAATAGCCGAGCCAGCCAGCCCTTCAAGAGCTGTCTTCCGCAAGTTCGATGAGCTTC 1380
QY 1178 CTTTATGAGCTTCACTTCAACATCACTGATGAAAGAGAGATGAGATTTCAAGCCAA 1237
Db 1178 CTTTATGAGCTTCACTTCAACATCACTGATGAAAGAGAGATGAGATTTCAAGCCAA 1237
QY 1381 CTTTATGAGCTTCACTTCAACATCACTGATGAAAGAGAGATGAGATTTCAAGCCAA 1440
Db 1381 CTTTATGAGCTTCACTTCAACATCACTGATGAAAGAGAGATGAGATTTCAAGCCAA 1440
QY 1238 TCAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1297
Db 1238 TCAGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1297
QY 1441 TCAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Db 1441 TCAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500

QY 1298 CCATGATCACCATCTCCCAAGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGA 1357
 DB 1501 CCAATGATCACCATCTCCCAAGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGA 1560
 QY 1358 ATCCCTTCTCCAGAGGCGCTCCCAAAAACCAAGGACCCAAACAGAACTTTAGGGG 1417
 DB 1561 ATCCCTTCTCCAGAGGCGCTCCCAAAAACCAAGGACCCAAACAGAACTTTAGGGG 1620
 QY 1418 CCAAGAAAGCAACAAGGCGTGAAGCTTAAAGCTGTGAGCCCTTCAAGCTTGCCCACT 1477
 DB 1621 CCAAGAAAGCAACAAGGCGTGAAGCTTAAAGCTGTGAGCCCTTCAAGCTTGCCCACT 1660
 QY 1478 GTATCAGAGGCGAGGCTACTACAGTGGCCCAAGAGCGCCCTCAGGCCCACTCCCATTT 1537
 DB 1681 GTATCAGAGGCGAGGCTACTACAGTGGCCCAAGAGCGCCCTCAGGCCCACTCCCATTT 1740
 QY 1538 CTTCCTCCCTAGAACATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGCTATGACACCA 1597
 DB 1741 CTTCCTCCCTAGAACATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGCTATGACACCA 1800
 QY 1598 AGACAAAAGCTTAAAGACTGTGAGTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGA 1657
 DB 1801 AGACAAAAGCTTAAAGACTGTGAGTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGA 1860
 QY 1658 GAGCGATGGGGCTTGAATGAGCAACCCAGAAATATCTCAAGTGAAGAGAAACCTGAGA 1717
 DB 1861 GAGCGATGGGGCTTGAATGAGCAACCCAGAAATATCTCAAGTGAAGAGAAACCTGAGA 1920
 QY 1718 GTTTTAACTGACGGATATGCCAGAGATCCCGAAATACCTCAAGAGACCTTTGGAACA 1777
 DB 1921 GTTTTAACTGACGGATATGCCAGAGATCCCGAAATACCTCAAGAGACCTTTGGAACA 1980
 QY 1778 ATCCCAACCAACATACACATCACTAAGATGATCACTGATCTTTTGGGCTTGGGA 1837
 DB 1981 ATCCCAACCAACATACACATCACTAAGATGATCACTGATCTTTTGGGCTTGGGA 2040
 QY 1838 AAAACAGGATGAGACATCTCTTAACCTGCTTCCATATGGGGATGCTTCCGCGACAGAGT 1897
 DB 2041 AAAACAGGATGAGACATCTCTTAACCTGCTTCCATATGGGGATGCTTCCGCGACAGAGT 2100
 QY 1898 CCTCAGCTGTGTGTACACACAGAGACACATGATCCAGTCAAGCCATACAGCTGTCCACA 1957
 DB 2101 CCTCAGCTGTGTGTACACACAGAGACACATGATCCAGTCAAGCCATACAGCTGTCCACA 2160
 QY 1958 CTGAAGAAAGCTGTCTTACACAGCTGATTAATGTTAGCTTATATGATTAATAAATCCC 2017
 DB 2161 CTGAAGAAAGCTGTCTTACACAGCTGATTAATGTTAGCTTATATGATTAATAAATCCC 2220
 QY 2018 AGACTACTTGCCTTTAAATGCCCTTTTATTCATATAAAAATCTGAAAGCTAGACGAACCA 2077
 DB 2221 AGACTACTTGCCTTTAAATGCCCTTTTATTCATATAAAAATCTGAAAGCTAGACGAACCA 2280
 QY 2078 TTGGAAACATTTAATCACTGAGTCTGAGTTCGAGTCCGGAACCTTATGTTCTATCTGAT 2137
 DB 2281 TTGGAAACATTTAATCACTGAGTCTGAGTTCGAGTCCGGAACCTTATGTTCTATCTGAT 2340
 QY 2138 CCAAGACAGCCACCTTAGTATCTGCGCAAACTAATGAGTTTATTAATTAATAAATCT 2197
 DB 2341 CCAAGACAGCCACCTTAGTATCTGCGCAAACTAATGAGTTTATTAATTAATAAATCT 2400
 QY 2198 CGTTAAAAAATTTTTTTTTT 2217
 DB 2401 CGTTAAAAAATTTTTTTTTT 2420

RESULT 8
 BC015220 2441 bp mRNA linear PRI 07-OCT-2003
 LOCUS
 DEFINITION Homo sapiens vitelliform macular dystrophy (Best disease, bestrorhin), mRNA (cDNA clone MGC:17761 IMAGE:3877806), complete cds.
 ACCESSION BC015220

VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 BC015220.2 GI:34783515
 MGC.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2441)
 Straube, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Hopkins, R.F., Stappert, B., Buell, K.H., Schaefer, C.F., Bhat, N.K., Diatchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L., Scheetz, T.E., Brownstein, M.J., Ueda, T.B., Tomihata, S., Carninci, P., Prange, C., Raha, S.S., Loughran, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Boeck, S.A., McEwan, P.J., McKernan, K.U., Malek, U.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hu, Y., S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 JOURNAL
 MEDLINE
 PUBMED
 12477932
 2 (bases 1 to 2441)
 Straube, R.
 Direct Submission
 Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Sep 16, 2003 this sequence version replaced gi:2195361.
 REMARK
 COMMENT
 Contract: MGC help desk
 Email: gcgds-remail.nih.gov
 Tissue Procurement: DCTD/DTF/Gazdar
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 14 Row: b Column: 9.
 Location/Qualifiers
 1..2441
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:17761 IMAGE:3877806"
 /tissue_type="Lung, large cell carcinoma"
 /clone_id="NIH_MGC_68"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..2441
 /gene="VMD2"
 /note="synonyms: BMD, BEST"
 /db_xref="locusid:7439"
 /db_xref="MIM:153700"
 573..2069
 /gene="VMD2"
 /codon_start=1
 /product="VMD2 protein"

/protein_id="AAH15220.1"
/db_xref="GI:22800661"
/db_xref="LocusID:7439"
/db_xref="MIT:153700"
/translation="MFEKLTLYCDSYIQLIPISFVLGFVYTLVITRMNOYENLPMD
RLMSLVSGFVSGKDEQGLJLRLLIRYANLGNVLISSVTAAYKRPSPQHLVOGFP
MTPAEHKLQELSLPHNMFVPMVPMANLMSKMLGRIIDPILLOSILNEMTLTQ
CGHLYADWMSIPLVTOVTVAVYSPFLICVIGROPIINRAKYPGHELDLVVETFE
LQPFYVGMLEKLSLAVDEMNDLPRMPPMYNKPPEPPPTAASQFRRASFWEST
FNISLKNEMEFOPNDEEDNAHGITGRVLQSDHDEHPRANSSRTKLMPKRESL
HEGLPKNHKAQKONVROGEBDKWKLAVAFKSAPLVQRGYSAPOPLSPTEPMFF
PLPSASKHSVTVGIDTQKSLKTVSSAKSKSPFLISBDGALMEHPVYSQVRKRTV
EFNLIDMPEIPEHNHKEPLBOSPLNTHITLKDHMDPYWALENDEAHS"
576..1427
/gene="VMD2"
/note="DUF289; Region: Putative membrane protein. This
family called family 8 in, may be a transmembrane protein
The specific function of this protein is unknown"
/db_xref="CD:pfam01062"

Query Match 80.4%; Score 1793; DB 9; Length 2441;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1887; Conservative 0; Mismatches 5; Indels 81; Gaps 1;

QY 257 GCTGGCCCTCAAGGAAACAACAGCTGATGTTGAGAACTGACTGTATTTGCGACG 316
DB 545 GCTGGCCCTCAAGGAAACAACAGCTGATGTTGAGAACTGACTGTATTTGCGACG 604
QY 317 CTACATCAGCTCATCCCATTTCCCTTCGTCGGCTTCTACAGTACGCTGCTGTAC 376
DB 605 CTACATCAGCTCATCCCATTTCCCTTCGTCGGCTTCTACAGTACGCTGCTGTAC 664
QY 377 CCCTGCTGGAACCAAGTACGAGAACCTGCGCTGCGCCGACCTCATGAGCCTGATC 436
DB 665 CCCTGCTGGAACCAAGTACGAGAACCTGCGCTGCGCCGACCTCATGAGCCTGATC 724
QY 437 GGGCTTCGTGGAAGGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 496
DB 725 GGGCTTCGTGGAAGGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 784
QY 497 CCGCAACCTGAGCAACGTCCTCATCTGCGACGCTGACGCGGAGCTACAGCCGCTT 556
DB 785 CCGCAACCTGAGCAACGTCCTCATCTGCGACGCTGACGCGGAGCTACAGCCGCTT 844
QY 557 CCGCAACCTGAGCAACGTCCTCATCTGCGACGCTGACGCGGAGCTACAGCCGCTT 616
DB 845 CCGCAACCTGAGCAACGTCCTCATCTGCGACGCTGACGCGGAGCTACAGCCGCTT 904
QY 617 GAGAAACCTGAGCCACACACATGTTCTGGGTGCGCGGTGCTGTTGGCAACT 676
DB 905 GAGAAACCTGAGCCACACACATGTTCTGGGTGCGCGGTGCTGTTGGCAACT 964
QY 677 GTCAATGAAGCGTGGCTTGAAGTCAATCCGGAACCTTACCTGCTCAGAGCCTGCT 736
DB 965 GTCAATGAAGCGTGGCTTGAAGTCAATCCGGAACCTTACCTGCTCAGAGCCTGCT 1024
QY 737 GAAAGATGAACAACCTTGAGTCACTCACTGAGCACTCTGATGAGCTTCTCTGAG 796
DB 1025 GAAAGATGAACAACCTTGAGTCACTCACTGAGCACTCTGATGAGCTTCTCTGAG 1084
QY 797 TATCCCACTGCTGATACACAGGTGATGAGTGGCGGTGATACAGCTTCTCTGAG 856
DB 1085 TATCCCACTGCTGATACACAGGTGATGAGTGGCGGTGATACAGCTTCTCTGAG 1144
QY 857 TCTAGTTGGGGGAGAGTTTCTGAACCTGAGCAAGCTTACCTGGCCATGAGCTGACCT 916
DB 1145 TCTAGTTGGGGGAGAGTTTCTGAACCTGAGCAAGCTTACCTGGCCATGAGCTGACCT 1204
QY 917 CGTTGGCCCGCTTTCAGCTTCTGAGAGTCTTCTTATGTTGAGTGGCTGAGAGTGC 976
DB 1205 CGTTGGCCCGCTTTCAGCTTCTGAGAGTCTTCTTATGTTGAGTGGCTGAGAGTGC 1257

QY 977 AGAGCAGCTCATCAACCCCTTTGGAGAGATGATGATGATTTTGAACCACTGATTC 1036
DB 1258 ----- 1257
QY 1037 CGACGAAATTTGACAGTGTGCTGTTGGCTGTGATGATGACCAAGACCTGCTCG 1096
DB 1258 ----- AGGTGTCCTGTTGGCTGTGATGATGATGACCAAGACCTGCTCG 1303
QY 1097 GATGAGCCGGAACATGATCTGGAATAAGCCCGGACCAAGCCCTTACACAGTGTTC 1156
DB 1304 GATGAGCCGGAACATGATCTGGAATAAGCCCGGACCAAGCCCTTACACAGTGTTC 1363
QY 1157 CGCCAGTTCCGTGAGCTTCCTTATGAGCTCCACCTTCAATCAGCTGAAACAAAG 1216
DB 1364 CGCCAGTTCCGTGAGCTTCCTTATGAGCTCCACCTTCAATCAGCTGAAACAAAG 1423
QY 1217 GGAAGTGAAGTTCCAGGCCCAATCAGAGGACGAGAGAGATGTCACGCTGGCATTTGG 1276
DB 1424 GGAAGTGAAGTTCCAGGCCCAATCAGAGGACGAGAGAGATGTCACGCTGGCATTTGG 1483
QY 1277 CCGCTTCTAGGCTGACATGCCATGATCACTCTCCAGGAGCAACCTCAAGAGCA 1336
DB 1484 CCGCTTCTAGGCTGACATGCCATGATCACTCTCCAGGAGCAACCTCAAGAGCA 1543
QY 1337 ACTACTGTGCCCAGAGAGGAATCCCTTCCACGAGGCTGCCCCAAACCAAGGC 1396
DB 1544 ACTACTGTGCCCAGAGAGGAATCCCTTCCACGAGGCTGCCCCAAACCAAGGC 1603
QY 1397 AGCCAAACAGAAAGTTAGGGGACAGAGAGACAAAGGCTGGAAGCTTAAAGCTGTGA 1456
DB 1604 AGCCAAACAGAAAGTTAGGGGACAGAGAGACAAAGGCTGGAAGCTTAAAGCTGTGA 1663
QY 1457 CGCTTCAAGTGTGCCCCACTGATACAGAGCCAGGCTTACTACAGTGGCCCCAGACGCC 1516
DB 1664 CGCTTCAAGTGTGCCCCACTGATACAGAGCCAGGCTTACTACAGTGGCCCCAGACGCC 1723
QY 1517 CCTCAGCCCACTCCCATGTTCTTCCCTTGAACCAATCAGGCGCTCAAGCTTCAAG 1576
DB 1724 CCTCAGCCCACTCCCATGTTCTTCCCTTGAACCAATCAGGCGCTTCAAGCTTCAAG 1783
QY 1577 TGTCAAGGCAATGACACCAAGACAAAGCTTAAAGCTGAGTTTGGGGGCAAGAA 1636
DB 1784 TGTCAAGGCAATGACACCAAGACAAAGCTTAAAGCTGAGTTTGGGGGCAAGAA 1843
QY 1637 AAGTTTGAATGCTCTGAGAGAGAGAGTGGGCTTGAATGAGACCCAGAAATATCTCA 1696
DB 1844 AAGTTTGAATGCTCTGAGAGAGAGAGTGGGCTTGAATGAGACCCAGAAATATCTCA 1903
QY 1697 AGTGAAGAGAAACTGTGAGTTTAACTGACGGAATGACAGAGATCCCGGAAATCA 1756
DB 1904 AGTGAAGAGAAACTGTGAGTTTAACTGACGGAATGACAGAGATCCCGGAAATCA 1963
QY 1757 CCTGAAACACTTTTGAACATACCAACCAATACACATACACTCAAAAGATCAT 1816
DB 1964 CCTGAAACACTTTTGAACATACCAACCAATACACATACACTCAAAAGATCAT 2023
QY 1817 GGAATCTTATGAGGCTTGAAGAACAGAGATGAAGCACTTCTAAGCTTCTTAATG 1876
DB 2024 GGAATCTTATGAGGCTTGAAGAACAGAGATGAAGCACTTCTAAGCTTCTTAATG 2083
QY 1877 GGAATGCTTGGCCAGCAGGCTCTGACCTGTGTGATACCCAGAGAGACATGATCCA 2143
DB 2084 GGAATGCTTGGCCAGCAGGCTCTGACCTGTGTGATACCCAGAGAGACATGATCCA 2196
QY 1937 ACAGCATAACAGCTGTGCAACATGGAAGAGAGTCTTCAACAGCTGGAATCAATGTT 1996
DB 2144 ACAGCATAACAGCTGTGCAACATGGAAGAGAGTCTTCAACAGCTGGAATCAATGTT 2203
QY 1997 AGCTTAATGAATAAAATCCAGACTACTTCAAGCTTATGAGCTTATTAATTAATAAC 2056
DB 2204 AGCTTAATGAATAAAATCCAGACTACTTCAAGCTTATGAGCTTATTAATTAATAAC 2263
QY 2057 TGTGAAGACTAAGACTAAGCACTTTAAGAACTTAAGTCAAGCTTCAAGAGTGGG 2116

Db	2264	TGTGAACCTAGACTGAAACCATTGGAAAACTTTAATCTAGACTCTGGATTCAAGTCCGG	2322
Oy	2117	AACCTTAGTCTATCTGATTCGAATCGAACAGCCACA CTTAGTATACTGCCCAAATAATG	2176
Db	2264	AACCTTAGTCTATCTGATTCGAATCGAACAGCCACA CTTAGTATACTGCCCAAATAATG	2383
Oy	2177	AGTTTAATTAATCAAAATATCTGCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2229
Db	2384	AGTTTAATTAATCAAAATATCTGCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2436
<hr/>			
RESULT 9			
LOCUS	AY151704	1758 bp	mRNA linear PRI 28-JAN-2004
DEFINITION	Homo sapiens bestrophin 1 mRNA, complete cds.		
ACCESSION	AY151704		
VERSION	AY151704.1	GI:41216872	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Sun,H., Teunenari,T., Yau,K.W. and Nathans,J. The vitelliform macular dystrophy protein defines a new family of chloride channels Proc. Natl. Acad. Sci. U.S.A. 99 (6), 4008-4013 (2002)		
JOURNAL MEDLINE	2 (bases 1 to 1758)		
PUBMED	21902086		
REFERENCE	Teunenari,T., Sun,H., Williams,J., Cahill,H., Smalwood,P., Yau,K.W. and Nathans,J. Structure-function analysis of the bestrophin family of anion channels J. Biol. Chem. 278 (42), 41114-41125 (2003)		
JOURNAL MEDLINE	22617481		
PUBMED	12907679		
REFERENCE	3 (bases 1 to 1758)		
AUTHORS	Sun,H., Teunenari,T., Yau,K.-W. and Nathans,J.		
JOURNAL TITLE	Submitted (02-JAN-2004) Molecular Biology, Neuroscience, Johns Hopkins University, 725 N. Wolfe Street, Baltimore, MD 21205, USA		
FEATURES	source		
	1..1758	/organism="Homo sapiens"	
		/mol_type="mRNA"	
	/db_xref="taxon:9606"		
CDS	1..1758	/note="Chloride channel"	
		/codon_start=1	
	/product="bestrophin 1"		
	/protein_id="AAR99654.1"		
	/db_xref="gi:41216873"		
	/translation="MTITVTSQVANARIIGSFSSRLLLCKRISITYKLLNGEPLIFLLCYW IIRFLRYLATLEEQDLMEPKLTLXCDSTYLIPISEFVGFTVLIVTRMNQYENLP PDRIMSLVSVEGEKDEQGRLLRTLYIANGLILRSVTAVYKSPSAOHLYOQ GFMPABHKOLEKLSLPHNMFVPWPVAFNISMKVAGGSRIDRPILIOSLNEMNTLAE TOCHLVAYDMISIPLYOVTVAYVVSFFETLCVGRQLPAAPARAYPEHELDIVPVIV TFLOPFYFGWLKVAEOILNPGEEDDFEFNNIYDRILOYSLAVIDMMDIDPRMEB DMWNKEPEPPPYPTNAASQFRRASWSSTNI SLNKEEMERQPNODEDEDHAAGLTGN PLGSASHDHHPRRANSRTKLMPKRESLLHGGLPNHNKAAKONVGOEDNKAMKLKA DAFGASAPLORGVYASAPOTPLSPTPMFPLEPSAPSXLHSVTDGTDKSLKTIVSSG AKSGELISESDGALMEHRVSYQRKTKVERNLTDMPEI PENHLKEPLEGSPTNHTTT LKQHNDPYWALENDRHNS"		
<hr/>			
ORIGIN			
Query Match	78.8%;	Score 1756.4;	DB 9; Length 1758;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 1757; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
<hr/>			
Oy	105	ATGACCATCCTTACACAAAGCCAATGGCTATGCCCGCTTAGGCTCTCTGCCGCTG	164

D	b	1	ATGACAACTTACACAGAGCAAGTGTGCTAATGCTCCGCTTAGCTCTTCTCCGCTG	60
O	y	165	CTGCTGTGCTGTGGCGGGGAGCATCTTACAGCTGCTGTATATGGCGAGTTCTTATTTCTCTG	224
D	b	61	CTGCTGTGCTGTGGCGGGGAGCATCTTACAGCTGCTGTATATGGCGAGTTCTTATTTCTCTG	120
O	y	225	CTGCTGTACTATCATCTCCGCTTATTTATATATGCTGGGCGCTCAGAGAAACAACAGCTG	284
D	b	121	CTGTGCTACTACATCATCTCCGCTTATTTATATATGCTGGGCGCTCAGAGAAACAACAGCTG	180
O	y	285	ATGTTTGAAGAACTGAGCTCTGTATTTGGGACAGCTACATCCAGCTCATCTCCCATTTCTTTC	344
D	b	181	ATGTTTGAAGAACTGAGCTCTGTATTTGGGACAGCTACATCCAGCTCATCTCCCATTTCTTTC	240
O	y	345	GTCCTGTGGCTTTCTACGTGACGCTGTGTGTGACCCGCTGTGTGGAAACAGTACGAACTGT	404
D	b	241	GTCCTGTGGCTTTCTACGTGACGCTGTGTGTGACCCGCTGTGTGGAAACAGTACGAACTGT	300
O	y	405	CCGTGTGGCCCAACCGGCTCATGAGGCTGTGTGTGTGGGCTTCGTCCGAAGGCAAGGAGACCA	464
D	b	301	CCGTGTGGCCCAACCGGCTCATGAGGCTGTGTGTGTGGGCTTCGTCCGAAGGCAAGGAGACCA	360
O	y	465	GGCGGGCTGTGCGGGGCGACGCTCATCTCCGCTACGCAACCTTGGGCAACGTTGCTCATCTGT	524
D	b	361	GGCGGGCTGTGCGGGGCGACGCTCATCTCCGCTACGCAACCTTGGGCAACGTTGCTCATCTGT	420
O	y	525	CGACGCGTCAAGCACCGGAGTCTTACAGCGCTTCCCGACGCGCCGACACTGTGTGCAAGCA	584
D	b	421	CGACGCGTCAAGCACCGGAGTCTTACAGCGCTTCCCGACGCGCCGACACTGTGTGCAAGCA	480
O	y	585	GGCTTTATGACTCCGGGAGAAACAACAAGCATTTGGAGAAACTGAGCCTTACCAACAACATGT	644
D	b	481	GGCTTTATGACTCCGGGAGAAACAACAAGCATTTGGAGAAACTGAGCCTTACCAACAACATGT	540
O	y	645	TTCTTGGGGGCGCTGTGGGTGTGTGTGGCCAACTGTGCAATGAAAGGTGTGAGTCTTGAGTCTGA	704
D	b	541	TTCTTGGGGGCGCTGTGGGTGTGTGTGGCCAACTGTGCAATGAAAGGTGTGAGTCTTGAGTCTGA	600
O	y	705	ATCCGGGAGCCCTTACTGTCTCCAGAGGCTGTGTAACGAGATGAAACACTTGTGCGTACTCAG	764
D	b	601	ATCCGGGAGCCCTTACTGTCTCCAGAGGCTGTGTAACGAGATGAAACACTTGTGCGTACTCAG	660
O	y	765	TGTGGAACACTGTATGCTTACGACTGTGGATTTATCCACTGTGTGTATACACAGTGTGTG	824
D	b	661	TGTGGAACACTGTATGCTTACGACTGTGGATTTATCCACTGTGTGTATACACAGTGTGTG	720
O	y	825	ACTGTGGGCGGTATCAGCTTCTTCTGTACATTTGTCTAGTTGGGCGGCACTTCTGAAACCA	884
D	b	721	ACTGTGGGCGGTATCAGCTTCTTCTGTACATTTGTCTAGTTGGGCGGCACTTCTGAAACCA	780
O	y	885	GCCAAAGGCTTACCTCTGGGCACTGAGCTGAGACTCGTTGTGCCGCTTTCACGTTTCTGCGAG	944
D	b	781	GCCAAAGGCTTACCTCTGGGCACTGAGCTGAGACTCGTTGTGCCGCTTTCACGTTTCTGCGAG	840
O	y	945	TTCTTTCTTCTATGTGTGCTGTGAGAGTGTGACAGGATTTGGCAGTGTCCCTGTGTG	1004
D	b	841	TTCTTTCTTCTATGTGTGCTGTGAGAGTGTGACAGGATTTGGCAGTGTCCCTGTGTG	900
O	y	1005	GATGATATATGATTTTGGAGCAACAATGTGATTTGTGACAGGAATTTGCAGAGTGTCCCTGTGTG	1064
D	b	901	GATGATATATGATTTTGGAGCAACAATGTGATTTGTGACAGGAATTTGCAGAGTGTCCCTGTGTG	960
O	y	1065	GCTGTGATGATGATGACACAGAGACTGTCTCGATGTGAGGCGGACATGTACTGTGAATTAAG	1124
D	b	961	GCTGTGATGATGATGACACAGAGACTGTCTCGATGTGAGGCGGACATGTACTGTGAATTAAG	1020
O	y	1125	CCGAGACCAAGCCCCCTTACACAGCTGTTCGCGCCAGTTCCGTGAGCTCTCTTTATG	1184
D	b	1021	CCGAGACCAAGCCCCCTTACACAGCTGTTCGCGCCAGTTCCGTGAGCTCTCTTTATG	1080
O	y	1185	GGCTTCAACTTTCAACATCAGCTGTGAACAAGAGAGATGTGATTCAGGCCAATCAGAGAG	1244
D	b	1081	GGCTTCAACTTTCAACATCAGCTGTGAACAAGAGAGATGTGATTCAGGCCAATCAGAGAG	1140

	Best Local Similarity	99.9%	Pred. No. 0;	Matches 1756;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	105	ATGACCACTACCTTACACAGCAAGTGGCTATAGCCCGCTTATAGGCTCTCTTCCCGCTG	164									
Db	1	ATGACCACTACCTTACACAGCAAGTGGCTATAGCCCGCTTATAGGCTCTCTTCCCGCTG	60									
QY	165	CTGCTGTCTGCGGGGACAGCATCTCAAGCTGTATATGGCGAGTTCTTATCTTCTG	224									
Db	61	CTGCTGTCTGCGGGGACAGCATCTCAAGCTGTATATGGCGAGTTCTTATCTTCTG	120									
QY	225	CTGTCTACTACATCATCCGCTTATTTATAGGCTGGGCTCCACAGGAACAACAGCTG	284									
Db	121	CTGTCTACTACATCATCCGCTTATTTATAGGCTGGGCTCCACAGGAACAACAGCTG	180									
QY	285	ATGTTTAAAGAACTGACTCTGTATTTGGACAGCTACATCCAGCTCATCCCATTTCTCTG	344									
Db	181	ATGTTTAAAGAACTGACTCTGTATTTGGACAGCTACATCCAGCTCATCCCATTTCTCTG	240									
QY	345	GTGTGGGGCTTTTACGTGACGCTGGTGTGACCCGCTGTGTGAACCACTGTACGAACCTG	404									
Db	241	GTGTGGGGCTTTTACGTGACGCTGGTGTGACCCGCTGTGTGAACCACTGTACGAACCTG	300									
QY	405	CCGTGCCCCGACCGCTCTACATGAGCTGTGTGTGGGCTCTGTCAAGGACAGACGACA	464									
Db	301	CCGTGCCCCGACCGCTCTACATGAGCTGTGTGTGGGCTCTGTCAAGGACAGACGACA	360									
QY	465	GGCCGGCTCTGCGCGACGCTCATCCGCTACGCAACCTGGGGCAAGTGTCTATCTTG	524									
Db	361	GGCCGGCTCTGCGCGACGCTCATCCGCTACGCAACCTGGGGCAAGTGTCTATCTTG	420									
QY	525	CGAGGTTACGACCGCAGTCTTACAGCGTCTCCAGCGCCAGACCTGGTGTGAACA	584									
Db	421	CGAGGTTACGACCGCAGTCTTACAGCGTCTCCAGCGCCAGACCTGGTGTGAACA	480									
QY	585	GGCTTTATGACTCCGCGACAAACAAGCTTGGAAACCTAGGCTTACACACAATG	644									
Db	481	GGCTTTATGACTCCGCGACAAACAAGCTTGGAAACCTAGGCTTACACACAATG	540									
QY	645	TTTCGGGTCCTTGGGTGTGTGCTTCCCACTGTCAATGAAGCCGTGCTTGGAGTCTGA	704									
Db	541	TTTCGGGTCCTTGGGTGTGTGCTTCCCACTGTCAATGAAGCCGTGCTTGGAGTCTGA	600									
QY	705	ATCCGGGACCTTATCTCGCTCCAGACCTGCTGGAACGAGTGAACACTTGTGCTGACG	764									
Db	601	ATCCGGGACCTTATCTCGCTCCAGACCTGCTGGAACGAGTGAACACTTGTGCTGACG	660									
QY	765	TGTGGAACCTGTATGTGCTTACGACTTGTAGTATCCCACTGTGTATACAGAGTGTG	824									
Db	661	TGTGGAACCTGTATGTGCTTACGACTTGTAGTATCCCACTGTGTATACAGAGTGTG	720									
QY	825	ACTGTGCGGTGACAGCTTCTTCCGACTTGTCAAGTGGGGCGAGTTCTTGAACCA	884									
Db	721	ACTGTGCGGTGACAGCTTCTTCCGACTTGTCAAGTGGGGCGAGTTCTTGAACCA	780									
QY	885	GCCAAAGGCTTACCTTGGCCATGAGCTGTGCTGTGCGCTTCTTACGTTCTTCCGAG	944									
Db	781	GCCAAAGGCTTACCTTGGCCATGAGCTGTGCTGTGCGCTTCTTACGTTCTTCCGAG	840									
QY	945	TTCTTCTTATAGTGTGGCTGGAAGGTGCAAGACGTCAATCAACCTTGTGAGAG	1004									
Db	841	TTCTTCTTATAGTGTGGCTGGAAGGTGCAAGACGTCAATCAACCTTGTGAGAG	900									
QY	1005	GATGATGATGATTTTGAACCACTGATTTGTGACAGGAATTTGAGAGTGTCCCTGTTG	1064									
Db	901	GATGATGATGATTTTGAACCACTGATTTGTGACAGGAATTTGAGAGTGTCCCTGTTG	960									
QY	1065	GCTGTGATGATGACACAGGACTTGTGCTGCTGATGAGACCGGACATGTATCTGAATTAAG	1124									
Db	961	GCTGTGATGATGACACAGGACTTGTGCTGCTGATGAGACCGGACATGTATCTGAATTAAG	1020									
QY	1125	CCGAGGCAACGCCCTTACAGACTTTCGCGCCAGATTCCGTGAGCTCTCTTATG	1184									

```

Db      1021 CCGGAGCCACAGCCCCCTACAGACGTGCTTCGCCGCCAGTTCCGTCGAGCTCCTTTATG 1080
Qy      1185 GGCTCCACCTTCAACATCAGCCTGAAACAAGAGGATGATGATTCAGGCCCAATCAGAG 1244
Db      1081 GGCTCCACCTTCAACATCAGCCTGAAACAAGAGGATGATGATTCAGGCCCAATCAGAG 1140
Qy      1245 GAGCAGAGAGATGCTCAACGCTGGATCAATGCGCGCTTCTTAGCGCTGACAGTCCATGAT 1304
Db      1141 GACGAGAGAGATGCTCAACGCTGGATCAATGCGCGCTTCTTAGCGCTGACAGTCCATGAT 1200
Qy      1305 CACCATCTCTCCAGAGGCAAACTCAAGACCAAACTACTGTGTGGCCCAAGAGGAATCCCTT 1364
Db      1201 CACCATCTCTCCAGAGGCAAACTCAAGACCAAACTACTGTGTGGCCCAAGAGGAATCCCTT 1260
Qy      1365 CTCACGAGGAGCTGCGCCAAACCAACAAGGAGCAAAAGCAAAAGTTAGGGGCGCAGGAA 1424
Db      1261 CTCACGAGGAGCTGCGCCAAACCAACAAGGAGCAAAAGCAAAAGTTAGGGGCGCAGGAA 1320
Qy      1425 GACAAACAAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGCTGTGCGCCACTGTATCAG 1484
Db      1321 GACAAACAAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGCTGTGCGCCACTGTATCAG 1380
Qy      1485 AGGCCAGGCTACTACAGTGTCCCAACAAGCGCCCTCAGGCCCACTCCCATGTTCTTCCC 1544
Db      1381 AGGCCAGGCTACTACAGTGTCCCAACAAGCGCCCTCAGGCCCACTCCCATGTTCTTCCC 1440
Qy      1545 CTAAACCATCAGGCGCGCTCAAAAGCTTCAAGCTGTGACAGGCAATGACACCAAGACAA 1604
Db      1441 CTAAACCATCAGGCGCGCTCAAAAGCTTCAAGCTGTGACAGGCAATGACACCAAGACAA 1500
Qy      1605 AGCTTAAAGACTGTAGTGTCTGGGCGCAAGAAAGTTGTAATGCTCTCAGAGAGCGAT 1664
Db      1501 AGCTTAAAGACTGTAGTGTCTGGGCGCAAGAAAGTTGTAATGCTCTCAGAGAGCGAT 1560
Qy      1665 GGGGCTTTGATGAGCACCAGAAAGTATCTTAAGTGAAGAGAAAGCTGTGAGTTTAC 1724
Db      1561 GGGGCTTTGATGAGCACCAGAAAGTATCTTAAGTGAAGAGAAAGCTGTGAGTTTAC 1620
Qy      1725 CTGACGATATGCGAGAGATCCCGGAAATACACCTTGAAGAACTTTGGAACATTCACCA 1784
Db      1621 CTGACGATATGCGAGAGATCCCGGAAATACACCTTGAAGAACTTTGGAACATTCACCA 1680
Qy      1785 ACCAATACACACTACACTCAAGATCAATGATCCTTATTTGGGCTTTGAAACAGG 1844
Db      1681 ACCAATACACACTACACTCAAGATCAATGATCCTTATTTGGGCTTTGAAACAGG 1740
Qy      1845 GATGAGCAGCATTCCTAA 1862
Db      1741 GATGAGCAGCATTCCTAA 1758

RESULT 11
BC041664 2435 bp mRNA linear PRI 07-OCT-2003
LOCUS BC041664 Homo sapiens vitelliform macular dystrophy (Best disease, bestrophin), mRNA (cDNA clone MGC:47884 IMAGE:5194649), complete cds
DEFINITION
ACCESSION BC041664
VERSION BC041664.1 GI:27371319
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2435)
AUTHORS Strauberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Helel,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schaefer,T.E., Brownstein,M.J., Udell,T.B., Tohbiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loggellano,N.A., Peters,G.J.,

```

```

TITLE
JOURNAL MEDLINE
PUBMED 22388257
REFERENCE 12477932
AUTHORS Strauberg,R.
TITLE Direct Submission
JOURNAL Submitted (20-DEC-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: mgc-help@nci.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Ahter,N., Ayala,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maseri,B., Mastrian,S.D., McCloskey,J.C.,
Madowell,J., Pearson,R., Stantrop,S., Thomas,P.J., Touchman,J.W.,
Tougeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAC Plate: 82 Row: d Column: 4.
FEATURES
SOURCE
location/Qualifiers
1. 2435
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:47884 IMAGE:5194649"
/tissue_type="Brain, adult, 6 pooled whole brains"
/clone_id="NIH_MGC_114"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1. 2435
/gene="VMD2"
/notes="synonyms: BMD, BEST"
/db_xref="LocusID:7439"
/db_xref="MIM:153700"
87..1901
/gene="VMD2"
/codon_start=1
/product="VMD2 protein"
/protein_id="AA41664.1"
/db_xref="GI:27371320"
/db_xref="LocusID:7439"
/db_xref="MIM:153700"
/translation="MPEKLTLYCDSTYGLIPSPVFGVYTLVTRMNOYENLPMPP
RLMSLVSGPREGDEGRRLRLTLRYANGLVILRSVTAAYKRPPSQHLYOAGF
MTAEKHOKLESLPNNMFVPPVWPAWSKMLGRIIDPILLOSLLNEKMLRTQ
CGELVYVDWISIPLVYTVQVTAIVYSEFLICAGROPLNPAKYVPGHEDLVVPPFTF
LGPFYVGMKVAEQLINPGEDEDDFETWIVDRNLQVSLAVDEKHQDLPRMEPDM

```

YMKRPEOPYTAASQFRRVSEWGFSTFNISLNKEEMEPQNEDEBDNAHGIISRL
GLQSHDHPHPRANSRTKLMPKRESLHLELPRNHKAKONNGQDNKAMKAYDA
FKSAPLYORPGYVSAPOPLSPMPPEPLPSAPSKLHSTGIDITOKSLKIVSGAK
KSPFLSESGALMEHPEVSOVRKRYERLUTWPELPEHMLKEPLVDSPTNHTLK
DHMDPYMALNRSYVHLNNOGCTALSLPFLNPLNFGHFCOSTLDRPAL
AMGIYLTFTFGLIKSCSGPFLTSPWHPEDFLDPSGR" 90.1022
/gene="VMD2"
/note="DUF289; Region: Putative membrane protein. This
family called family 8 in. may be a transmembrane protein
The specific function of this protein is unknown"
/db_xref="CD:pfam01062"

ORIGIN

Query Match 71.1%; Score 1584; DB 9; Length 2435;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1587; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 257 GCTGGCCCTCAGGAGAAACAAGCTGATGTTTGAAGAACTGACTGTGTATTGGCAGAG 316
Db 59 GCTGGCCCTCAGGAGAAACAAGCTGATGTTTGAAGAACTGACTGTGTATTGGCAGAG 118
QY 317 CTACATCAGCTCATCCCATTTCTTCGNGCTGGGCTTCTAGTGAAGCGTGTGTAC 376
Db 119 CTACATCAGCTCATCCCATTTCTTCGNGCTGGGCTTCTAGTGAAGCGTGTGTAC 178
QY 377 CCGCTGTGGAAACAGTACGAGAACCTGCGTGGCCGACCGCTCATAGCGTGTGTAC 436
Db 179 CCGCTGTGGAAACAGTACGAGAACCTGCGTGGCCGACCGCTCATAGCGTGTGTAC 238
QY 437 GGGCTTCGTGAAGGAG 496
Db 239 GGGCTTCGTGAAGGAG 298
QY 497 CCGCAACTGGGCAACGTGCTCATCTCGGCAAGCGTCAAGACCGGAGTCAAGAGCTT 556
Db 299 CCGCAACTGGGCAACGTGCTCATCTCGGCAAGCGTCAAGACCGGAGTCAAGAGCTT 358
QY 557 CCGGAGCGCCAG 616
Db 359 CCGGAGCGCCAG 418
QY 617 GAGAAACTGAGAGCTTACCAACAATGTTCTGGGTGCGGCTGAGTGTGCTTGGCAACT 676
Db 419 GAGAAACTGAGAGCTTACCAACAATGTTCTGGGTGCGGCTGAGTGTGCTTGGCAACT 478
QY 677 GTCAATGAGAGCGTGGCTTGGAGGTGCAATCGGAGCTTATCTGCTCCAGAGCTGCT 736
Db 479 GTCAATGAGAGCGTGGCTTGGAGGTGCAATCGGAGCTTATCTGCTCCAGAGCTGCT 538
QY 737 GAAAGAGTGAACAACCTTGGCTGCTAGTGTGAGACCTGTATGCTTATGCTTGAAG 796
Db 539 GAAAGAGTGAACAACCTTGGCTGCTAGTGTGAGACCTGTATGCTTATGCTTGAAG 598
QY 797 TATCCCACTGTGTATACAGAGTGTGACTGTGGCGGTGTAGAGCTTCTCTGACTTG 856
Db 599 TATCCCACTGTGTATACAGAGTGTGACTGTGGCGGTGTAGAGCTTCTCTGACTTG 658
QY 857 TCTAGTGGGCGGAGTGTCTGAACCCAGGCAAGGCTTACCTGCGCATGAGTGAAGCTT 916
Db 659 TCTAGTGGGCGGAGTGTCTGAACCCAGGCAAGGCTTACCTGCGCATGAGTGAAGCTT 718
QY 917 CGTTGTGCGGCTTTCAGTTCCTGAGTGTCTTCTATGTTGGCTGCTAAGGTGCG 976
Db 719 CGTTGTGCGGCTTTCAGTTCCTGAGTGTCTTCTATGTTGGCTGCTAAGGTGCG 778
QY 977 AGAGAGCTATTAACCCCTTTGAGAGAGTATGATATTTGAGACCAACTGATGTT 1036
Db 779 AGAGAGCTATTAACCCCTTTGAGAGAGTATGATATTTGAGACCAACTGATGTT 838
QY 1037 CGACAGAAATTTGAGAGTGTCTGTTGGCTGAGTGAAGAGCAACGAGAGCTGCTCG 1096
Db 839 CGACAGAAATTTGAGAGTGTCTGTTGGCTGAGTGAAGAGCAACGAGAGCTGCTCG 898

QY 1097 GATGAGCCGAGACATGTAATGAAATTAAGCCGAGCCACAGCCCCCTACACAGTGTTC 1156
Db 899 GATGAGCCGAGACATGTAATGAAATTAAGCCGAGCCACAGCCCCCTACACAGTGTTC 958
QY 1157 CGCCCAATTCGTGAGCTCTCTTATGAGCTCCACCTTCAATCAGCTGGAACAAGA 1216
Db 959 CGCCCAATTCGTGAGCTCTCTTATGAGCTCCACCTTCAATCAGCTGGAACAAGA 1018
QY 1217 GGAATGAGAGTTCAGCCCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1276
Db 1019 GGAATGAGAGTTCAGCCCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1078
QY 1277 CCGCTTCTTAGGCTGAGTCCATGATCAATCTCTCCAGGCAAGCTTAAGAGCAAA 1336
Db 1079 CCGCTTCTTAGGCTGAGTCCATGATCAATCTCTCCAGGCAAGCTTAAGAGCAAA 1138
QY 1337 ACTACTGAGCCCAAGAGAGAGATCCCTTCTCCAGAGGCTTGCCTCAAAATCAAGAGC 1396
Db 1139 ACTACTGAGCCCAAGAGAGATCCCTTCTCCAGAGGCTTGCCTCAAAATCAAGAGC 1198
QY 1397 AGCCAAACAGAAAGTTAGGGGCCAGAGAGACAAAGAGCTGAAAGCTTAAGCTGTGA 1456
Db 1199 AGCCAAACAGAAAGTTAGGGGCCAGAGAGACAAAGAGCTGAAAGCTTAAGCTGTGA 1258
QY 1457 CGCCTTCAAGTCTGGCCCACTGATATGAGAGCCAGGCTTACAGTGTGCCCAAGAGCC 1516
Db 1259 CGCCTTCAAGTCTGGCCCACTGATATGAGAGCCAGGCTTACAGTGTGCCCAAGAGCC 1318
QY 1517 CCTCAGCCCACTCCCATATTTCTTCCCTTAGAACATCAGAGGCGGTCAAAAGCTTCAAG 1576
Db 1319 CCTCAGCCCACTCCCATATTTCTTCCCTTAGAACATCAGAGGCGGTCAAAAGCTTCAAG 1378
QY 1577 TGTCAAGGATAGACACCAAGACAAAGCTTAAAGCTGATGTTTGGGGCCAAAGAA 1636
Db 1379 TGTCAAGGATAGACACCAAGACAAAGCTTAAAGCTGATGTTTGGGGCCAAAGAA 1438
QY 1637 AAGTTTGAATTCCTCAG 1696
Db 1439 AAGTTTGAATTCCTCAG 1498
QY 1697 AGTGAAGAGAAACTGTGAGATTAACTGACGATATGACAGATATCCCGAAAAATCA 1756
Db 1499 AGTGAAGAGAAACTGTGAGATTAACTGACGATATGACAGATATCCCGAAAAATCA 1558
QY 1757 CCTCAAGAACCTTGGAGAACATCAACCAACATACATACATCACTCAAAAGTCAAT 1816
Db 1559 CCTCAAGAACCTTGGAGAACATCAACCAACATACATACATCACTCAAAAGTCAAT 1618
QY 1817 GGATCCTTATTTGGGCTTGGAAAAACAGGAGATG 1848
Db 1619 GGATCCTTATTTGGGCTTGGAAAAACAGGATCTG 1650

RESULT 12
BD132659 1263 bp DNA linear PAT 18-SEP-2002
LOCUS BD132659
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD132659
VERSION BD132659.1 GI:23227604
KEYWORDS JP 2002504822-A.2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1263)
JACOBS K., MCCOY J.M., LAVALLIE E.R., RACIE J.A., TREAC M.,
SPAUDLING V., AGOSTINO M.J., HOWES S.H. and FECHTEL K.
Secreted proteins and polynucleotides encoding them
JOURNAL JOURNAL
GENETICS INSTITUTE INC
PN JP 2002504822-A/2
PD 12-FEB-2002

[illegible]

Db	781	CTCAAGAACCTTTGGAAACATACCAACCAACTACACTACACTCAAAAGATCACATG	840
Qy	1818	GATCCTTATTGGGCTTTGGAAAAACAGGATGAGACACATTCCTAACCTGCTTCTTAATGG	18777
Db	841	GATCCTTATTGGGCTTTGGAAAAACAGGATGAGACACATTCCTAACCTGCTTCTTAATGG	900
Qy	1878	GGATGCTTCGCGACCCAGGCTCCACTGCTGTGTATACACCCAGACAGACATGATCCAGCA	1937
Qy	1938	CAGCCATACAGCTGTCCACACTGAAGAAGCTGTCTCAACAGCGCTGATCAAAATGGTTA	1997
Db	961	CAGCCATACAGCTGTCCACACTGAAGAAGCTGTCTCAACAGCGCTGATCAAAATGGTTA	10200
Qy	1998	GCTTAATAGATAAAAATCCAGACTACTTCAGCCTTTAATGCTTTATTATCAAAAACT	20577
Db	1021	GCTTAATAGATAAAAATCCAGACTACTTCAGCCTTTAATGCTTTATTATCAAAAACT	10800
Qy	2058	GTGAAGACTAGACTGGAACCATTTGAAACATTTAATCTAGACTCTGAGATTACAGATCCGGGA	2117
Db	1081	GTGAAGACTAGACTGGAACCATTTGAAACATTTAATCTAGACTCTGAGATTACAGATCCGGGA	11400
Qy	2118	ACCTTGTGCTTATCTGTGATTCGAACAGACAGCCACCTTATGATATCTGCCCAACTAATGA	21777
Db	1141	ACCTTGTGCTTATCTGTGATTCGAACAGACAGCCACCTTATGATATCTGCCCAACTAATGA	12000
Qy	2178	GTTTAATTAATCAAAATACCTCGTTAAAAA	2229
Db	1201	GTTTAATTAATCAAAATACCTCGTTAAAAA	1252
RESULT 13			
BC079048		2035 bp	linear
LOCUS	Rattus norvegicus	CDNA	clone MGC:93985 IMAGE:7115295, complete cds.
DEFINITION			
ACCESSION	BC079048		
VERSION	BC079048.1	GI:50927692	
KEYWORDS			
SOURCE			
ORGANISM			
	Rattus norvegicus (Norway rat)		
	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE			
AUTHORS	1 (bases 1 to 2035)		
	Straubergs,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,		
	Klausner,R.D., Collins,F.S., Wagner,L., Schaffer,C.M., Schuler,G.D.,		
	Altschul,S.F., Zeeberg,B., Buelow,K.H., Sherman,C.F., Bhat,N.K.,		
	Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,		
	Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,		
	Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,		
	Scieczek,T.E., Brownstein,M.J., Uedlin,T.B., Tooby,N.K., S.		
	Carninci,P., Prange,C., Raha,S.S., Loggellano,A., Peters,G.J.,		
	Abrahamson,R.D., Mullaly,S.J., Boeck,S.A., McEwan,P.J.,		
	McKernan,K.D., Malek,U.A., Gunaratne,P.H., Richards,S.,		
	Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,		
	Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,		
	Fahney,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,		
	Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,		
	Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,		
	Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,		
	Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,		
	Schmeckel,A., Schein,J.E., Jones,S.J., and Marz,M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length		
	human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PIRATED	12477932		
REFERENCE	2 (bases 1 to 2035)		
AUTHORS	Director MGC Project.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-AUG-2004) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		

Qy 1487 GCCAGCTACTACAGTGGCCCCACAGAGCCCTCGACCTCCAGTGTCTTCCCTC 1546
Db 1511 GCCAAGTTACCGCCAGGCCCCACAGCCACTTGGCAGCCACTTCCAGTGTCTTCCCTC 1570

Qy 1547 AGAACCATCAGCGCCGCTCAAGCTTCAAGTGTCAAG 1585
Db 1571 TCAGCAGTCAGCAGCCCTCAGTGCCTTACTCAGACACAG 1609

RESULT 14
AY450427 1904 bp mRNA linear ROD 24-JUN-2004
LOCUS Mus musculus vitelliform macular dystrophy 2 (Vmd2) mRNA, complete cds.
DEFINITION
ACCESSION AY450427 GI:41612988
VERSION AY450427.1
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1904)
Stoehr, H., Marquardt, A., Nanda, I., Schmid, M. and Weber, B. H.
Three novel human VMD2-like genes are members of the evolutionary highly conserved RFP-TM family
Eur. J. Hum. Genet. 10 (4), 281-284 (2002)
JOURNAL MEDLINE 22027749
PUBMED 12032738
REFERENCE 2 (bases 1 to 1904)
Kraemer, F., Stoehr, H. and Weber, B. H. F.
Cloning and characterization of the murine Vmd2 RFP-TM gene family
Cytogenet. Genome Res. 105 (1), 107-114 (2004)
JOURNAL 3 (bases 1 to 1904)
Weber, B. H. F. and Kraemer, F.
Direct Submision
Submitted (27-OCT-2003) Human Genetics, University of Wuerzburg, Am Hubland, Bioentre, Wuerzburg 97074, Germany
FEATURES
SOURCE
location/Qualifiers
1..1904
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="19"
/cbase_type="teecis"
1..1904
/gene="Vmd2"
37..1692
/gene="Vmd2"
/codon_start=1
/product="vitelliform macular dystrophy 2"
/protein_id="AA509922.1"
/db_xref="GI:41612989"
/translation="MTITYNKVNANRLGSPSSLLCMRGSYKLYGEFLVFIPLY
SIRGLYRMVLSSDQQLPEKALYCDSTYDILPISFYLQVPTLVSRMSOYENLPW
PDRIMIOVSFVEKDEEGRILRLIRYALIGOVILIRISITVYRFTLHLVLA
GFMTGHEKQIQKGLPHNTFWVFWFANLSMAYIGRIRDTVLQSLMNEVCTLR
TOGQLAYDMISIPLYYQVVTVAVSFLACILIGQPLPNKDYGHEMDLVVPF
TILQFLPYMGLKVAEQILNPFGEDDDDPETNMIIDRLQVSLSDGMHONLPMER
DMYNBNAPOBPYPAASRRBSRBSFMGSTFNISLKEDLELMSKEADTDKESGVSS
TTCGFTGLQIKNYHLPLKDLKTKYLCSKNPLRSGCDANQKDKWKRKGLDPLKC
VPRKRGSGCGPAPSSHPTEBSAPSSSTGDSFSTDYQETCHMKKTVBFNLTPE
SPTEHLQORRLDQNSTNIOALMKENASYPYRDEAGTKPYLYE"

ORIGIN
Query Match 40.1%; Score 893.4; DB 10; Length 1904;
Best Local Similarity 77.3%; Pred. No. 6.5e-191;
Matches 1140; Conservative 0; Mismatches 236; Indels 39; Gaps 3;

Qy 69 GACCAAGCCCACTGTGACGCCCACTGCTGGCCATGACCATCACTTACAAAGCAA 128
Db 1 GACCAAGCCCACTGTGCTGCGCCAGTGCACCATGATCATCACTACAAACAAA 60

Qy 129 GTGCTATGCGCCGCTTAAAGCTCTTCCCGCTGCTGTGTGTGAGGAGCAGATC 188
Db 61 GTAGCCATGCGCGCTCGGTTGCTTCTGCTCTCTCTGCTGCTGCGAGGAGCATC 120

Qy 189 TACAAGCTGTATATGCGAGTTCTTATCTTCTGCTGTGCTATCAATCATCCGCTT 248
Db 121 TACAAGCTGTATATGAGAAATTCCTTGTCTTCAATATCTCTATATTCATCGGTGA 180

Qy 249 ATTTATAGCTGCGCCCTCAGGAAAGAACACAGCTGATGTTTGAAATCACTCTGAT 308
Db 181 CTTCACGAATGATGTTCTCGAGTATCAGCAGCTGTGTTGAAAGCTGCTGTGAC 240

Qy 309 TGCACAGCTACATCCAGCTCATCCCATTTCTCTGCTGCTGCTCTTACGTAGCGTG 368
Db 241 TGCACAGCTACATCCAGCTCATCTTATCTCTGCTGCTGCTCTTATGTTACATTC 300

Qy 369 GTCTGACCCGCTGTGGAACCAATACAGAACCTGCTGAGCCGACCGCTCATGAGC 428
Db 301 GTGCTGAGCCGCTGTGAGCAGTACAGAACTTGCCTGAGCCGACCGCTCATGATC 360

Qy 429 CTGCTGCGGCTTCTGTCGAAGCAGAGAGCAGAGCGGCTGCTGCGGCGACGCTC 488
Db 361 CAGGTGCTAGCTTCTGTCGAAGGAGCAGAGTGAAGAGCGCTTGTGCTGCGGCGACGCTC 420

Qy 489 ATCCGCTACGCGCAACCTGAGGAGAGCTGCTATCTGCGAGCGCTACACCGAGCTTAC 548
Db 421 ATCCGCTACGCGCAACCTGAGGAGAGCTGCTATCTGCGAGCATACAGCATCTCGCTTAC 480

Qy 549 AAGCGCTTCCCGACGCGCCAGCAGCTGTGTCAGACAGAGCTTATGATCTCGGAGAAAC 608
Db 481 AAGCGCTTCCCGACTTCTTACACACCTGCTGTACAGAGTTTATGACCAATGGGAAAT 540

Qy 609 AAGCATTTGAGAAATCAGAGCTTACCAACAACATGTTCTGAGTGCCTGAGTGTGTT 668
Db 541 AAGCATTTGAGAAATGAGGCTTACCAACAACATGTTCTGAGTGCCTGAGTGTGTT 600

Qy 669 GCCAAGCTGCAATGAAGAGGCTGAGTGTGAGTGAATCCGAGAGCTTATCTGCTCAG 728
Db 601 GCCAAGCTGCAATGAAGAGGCTTATCTGAGAGTGAATCCGAGAGCTTATCTGCTCAG 660

Qy 729 AGCTGTGACGAGATGAACACCTTGTGCTGATCTGATGAGACCTGTATGCTTACGAC 788
Db 661 AGCTGTGATGATGAGTGTGATCTTGTGCTGATCTGATGAGTGTGATGCTTACGAC 720

Qy 789 TGGATTATGATCCCATGCTGTGTATACACAGGTGTGATCTGCGGTGTACAGCTTCTC 848
Db 721 TGGATTATGATCCCATGCTGTGTATACACAGGTGTGATCTGCGGTGTACAGCTTCTC 780

Qy 849 CTGATCTGTCTAGTTGGGCGGAGTTCGTGAACCCAGACCAAGGCTTACCTGGGCAATGAG 908
Db 781 CTGATCTGTCTAGTTGGGCGGAGTTCGTGAACCCAGACCAAGGCTTACCTGGGCAATGAG 840

Qy 909 CTGACCTGCTGTGACCGCTTCTACAGCTTCTGCTGAGTTCCTTATGTTGAGTGTGCTG 968
Db 841 ATGATATGCTGTGCTGTCTTCTTACATCTGTGAATTTCTTATTTCAATGGGCTGCTG 900

Qy 969 AAGTGGCAAGCAGCTCATCAACCCCTTGGAGAGATGATGATTTTGAACCAAC 1028
Db 901 AAGTGGCAAGCAGCTCATCAACCCCTTGGAGAGATGATGATTTTGAACCAAC 960

Qy 1029 TGGATTGTGACAGGAATTTTGGAGGTTCCTCTGTTGCTGTGATGATGACACAGAGAC 1088
Db 961 TGGATTGTGACAGGAATTTTGGAGGTTCCTCTGTTGCTGTGATGATGACACAGAGAC 1020

Qy 1089 CTGCTGTGATGAGCGGAGCATGTATGGAATTAAGCCGACAGCCCTTACACA 1148
Db 1021 TTGCTCTCCATGAGACGTGACATGTATGGAACGAGGACAGCGCTTACACA 1080

Qy 1149 GCTGCTTCCGCGCAATTCCTGAGAGCTCTTATAGGCTTCCACTTCAACATGAGCTG 1208
Db 1081 GCTGCTTCCGCGCAATTCCTGAGAGCTCTTATAGGCTTCCACTTCAACATGAGCTT 1140

QY	1209	AACCAAGGAGATATGAGCTTCCAGGCCCAATCAGAGGACGAGAGAGATGCTACGCTG--	1266
Db	1141	AGGAAAGAAAGCTTGAAGCTTTGGTCAAAAAGGAGGAGCTGACACCGATTAAAGAAAGAGT	1200
QY	1267	-----GATCATGTCGCCGCTCTCTAGAGCTCGAGTCCCATGATCACCATGCTCCG	1314
Db	1201	GGCTATTAGCAGACCATATAGGCTGTCTTTTAAAGACTGCAACCCAAAACTACCATCTTCC	1266
QY	1317	AGGGCAAACTCAAGGACCAAAATACATGTCGCCAAGAGGAATCCCTTCTCAGAGAGGC	1378
Db	1261	TTGAAGAGCTTTAAAGACCAAACTATTGGTTTAAAG-----ATCCCTCTCTGAAAGGC	1314
QY	1377	CTGCCAAAAACCAAGGACGACCCAAACAGAACTTTAAGGGCCAGAAAGACAAACAAGGC	1438
Db	1315	CAGTGTAAGGATGCCCAACCCAGAAAAACAGAAAG-----ATGTC	1353
QY	1437	TGGAAGCTTAAAGGCTGTGAGACGCTTCAAGTCTGGCCCATGATATCAGAGCCAGGCTAC	1438
Db	1354	TGGAATTTAAAGGCTGTGAGCTTCTTGAAATATGTGTTCCAAAGTTTAAAGAGAGAGGCTCC	1413
QY	1497	TACAGTCCCCCAAGAGCGCCCTTACGCCCACTCC	1531
Db	1414	CATTGTGGCCCAAGGACCAACGAGCCACCTTAC	1448

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
15	BD136744	Beet's macular dystrophy gene.	BD136744.1	GI:3231689	JF_2002504559-A/26.	Mus musculus (house mouse)
	BD136744	1916 bp				
		DNA				
		linear				
		PAT 18-SEP-2002				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 1916)	Petrushin, K., Caskey, T.C., Metzker, M. and Wadelius, C.	Beet's macular dystrophy gene	Patent: JP 2002504559-A 26 12-FEB-2002;	
	MERCK & CO INC., CLASB WADELINUS			
	OS Mus musculus (mouse)			

FEATURES	Location/Qualifiers	Organization	Source
1. .1916		mus musculus (mouse)	

ORIGIN

Query Match	39.1%	Score 871.2;	DB 6;	Length 1916;
Best Local Similarity	77.1%;	Pred. No. 6.8e-186;		
Matches 116;	Conservative	0;	Mismatches 293;	Indels 39;
				Gaps 3;

QY 96 TGGCTGGGCATGACCATTCCTTACAAAGCAGGTGGCTATATGCCCGCTTAGAGCTCTTC 155

Db 2 TGGCAAGCATGACTCTACCTTACCAAAAGATGAGCCAAATGCCCGCTCGATTGTC 61

QY 156 TCCCGCTGTGCTGTGTGTGGCGGGGAGCATTTACAAGCTGTGTATATGCGAGTTCCTTA 215

Db 62 TGTGTCCTCTCTGTGTGTGGCGGAGCATTTACAAGCTGTGTATGAGAAATTCCTT 121

Qy	216	ATTTCTCTGCTGCTGCTACTACATCAATCCGCTTTATTTATTAAGCTGGCCCTCACGGAAGAA	275
Db	122	GTCTTATATATTCCTCTACTACTATTCATCCGCTGGACTCTCAAGATGGTCTCTCGATGTAT	181
Qy	276	CACAGCTATGTTTATAGAAACTGATCTGTATTTGGAACAGTCAATCCAGCTCATCCCC	335
Db	182	CAGCAGCTTGTTGTTTATAGAAAGCTGGCTGTACTGTGGACAGCTACATTCAGCTCATCCCT	241
Qy	336	ATTTCCTTGTGTGTGGGCTTCTACGTGACGCTGTGTCTGACCCGCTGTGTGAACCAATAC	395
Db	242	ATATCTTGTGTTCTGGGTTTCTATGTATATTGTGTGTAGCCGCTGTGTGAAGCCAGTAC	301
Qy	396	GAGAACTTCCGCTGGGCCACAGCCCTCATGAGCTGTGTGTGGGGCTTGTCTGAAGGCAAG	455
Db	302	GAGAACTTGCCTGTGGCCGACCCGCTCATGATTCAGAGTGTCTAGCTTCTGTGAAGGGCAAG	361
Qy	456	GACGAGCAAGGCTCGGCTGTGTGTGGCGACGCTCATTCGCTACGCCAACCTGTGGCAAGTG	515
Db	362	GATGAGGAAGGCGCTTTGTGTGTGGGGGACGCTCATTCGCTACGCCATCTGTGGGCCAAGTG	421
Qy	516	CTCATCTGTGCGAGGCTGACGACCGCAGCTTCAAGAGCTTCCCGACGSCCAGAGCCAGACTG	575
Db	422	CTCATCTGTGCGAGATCAGACCTCGGTCTTCAAGAGCGCTTTCCACTCTTCAACACCTGG	481
Qy	576	GTGCAAGCAAGCTTTATATGATCTCCGGCAGAAACAAGAGTTGTGAAGAACTGACCTTACCA	635
Db	482	GTGCTAGGAGGTTTATATGATCCCATGTGGGAAACAATAGCAGTTGTCAAGATGTGGGCTTACA	541
Qy	636	CACAACTATGTTCTGGGTGCCCCGTGGGTGTGTGTGCAACCTGTCAATAAAGCCGTGGCT	695
Db	542	CACAACTATGTTGTGGGTGCCCCGTGGGTGTGTGTGCAACCTGTCAATAAAGCCCTTACTT	601
Qy	696	GGAGGTGAATCCGGGACCTTATCTGTCTCAGAGCCCTGTGAACGAGATGAACACTGTG	755
Db	602	GGAGGTGCAATCCGGGACACCCTCTGTCTCAGAGCCCTGATGAATGAAGTGTGATCTTTG	661
Qy	756	GGTACTGAGTGTGACACTCTGTATAGCCCTTACGACTGTGAATTAATATCCCACTGTGTATACA	815
Db	662	GGTACTGAGTGTGTGACAGCTGTATGTCTTACGACTGTGAATTAATATCCCACTGTGTATACA	721
Qy	816	CAGGTGTGACTGTGTGCGCGGTGTACAGCTTCTTCCGATCTTGTATAGTTGTGGCGGAGTTT	875
Db	722	CAGGTGTGTGACAGTGTGCGCGGTGTACAGCTTCTTCCGATCTTGTATAGTTGTGGAGCGAGTTT	781
Qy	876	CTGAACCCAGCCAGAGGCTTACCTCTGGCCATGAGCTGGACTCTGTGTGTGCCGTCTTACG	935
Db	782	CTGAACCCAAACAAGGACTACCCAGGCGCATGATGTATCTGTGTGTGCCGTCTTACACA	841
Qy	936	TTCTGTGAGTCTTCTTCTATGTGTGGCTGTGGGTGAAGGTGGCAGAGAGCTATCAACCCC	995
Db	842	ATCTGTGAATTTCTTATTTCTACATGGGCTGTGTGAAGGTGGCAGAACAGCTATCAACCCC	901
Qy	996	TTTGTGAAGATGATGATGATTTTGAACCAACTGTGATTTGTGACAGAGATTTGCAGGTG	1055
Db	902	TTGTGGGAGGACGATGATGATTTTGAACACTTACGTGATCTATTTGACAAAAACTCTCAAGTG	961
Qy	1056	TTCCTGTGTGCTGTGATGATGATTCACACAGGACCTGTGCTGGATTTGAGGCGCGGACATGTAC	1115
Db	962	TTCCTGTGTGTGCGTGTGATGTGATGTACACAAACTTGTGCTCCATGTAACTGTACATGTAC	1021
Qy	1116	TGGAATATAGCCCGAGCCACAGCCGCCCTTACACAGAGCTTTCGCGCCAGTTTCGTGAGCC	1175
Db	1022	TGGAATATAGGCGAGGCTCTCAGCGCCCTTACACAGAGCTTCTTGTCCAGGTCTCGCGGCAT	1081
Qy	1176	TTCCTTATGTGGGCTGCACCTTCAACATCAGGCTGAACAAAGAGAGATGAGATTCCAGCCC	1235
Db	1082	TTCCTTATGTGGGCTGCACCTTCAACATCAGGCTTAAAGAGAAAGACTTATGAGCTTTGTGTCA	1141
Qy	1236	AATCAGAGGACGAGAGAGATGTCTACGCTG-----GCATCATTTGGCGCGCTTC	1283
Db	1142	AAAGAGGAGGCTGACACGATATAGAAAGAGAGTGGGCTATAGCAGACCAATAGGCTGTCTC	1201

```

QY 1284 CTAGGCTGCGATGCCATGATCAACATCTTCCAGGGCAAACTCAAGACCAACTACTG 1343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1202 TTAGGACTGCACCAAAAATACTATCTTCCCTTGAAAGACTTAAAGACCAAACTATTG 1261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1344 TGGCCCAAGAGGGAATCCCTTCTTCAGAGGGCTGCCCCAAAACCAAGGCAAGCCAA 1403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1262 TGTTCCTAAGA-----ACCCCTCTCTCGAAGGCCAGTAAAGATGCCAAACAGAAAAAC 1315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1404 CAGAACGTTAGGGGCCAGAAAGACAAGGCTGGAAGCTTAAAGGCTGTGAGCGCCTTC 1463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1316 CAGAAAG-----ATGCTGGAATTTAAGGCTCTGACTTCTTG 1354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1464 AAGTCTGCCCCACTGTATCAGAGGCCAGGCTACTACAGTCCCCACAGACCCCTCAGC 1523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1355 AATGTGTTCCAAAGTTTAAAGAGAGAGGCTCCCATTTGTGCCCCACAGGCAACCAAGCAGC 1414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1524 CCACTTCC 1531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1415 CACCTTAC 1422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: March 26, 2005, 07:38:20
 Job time : 9510.76 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 22:43:02 ; Search time 7168.13 Seconds
(without alignments)
11836.456 Million cell updates/sec

Title: US-09-622-964a-2
Perfect score: 2229
Sequence: 1 caggagagctccaccagccta.....aaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479086

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : EST.*
1: gb_eest1.*
2: gb_eest2.*
3: gb_hc1.*
4: gb_eest3.*
5: gb_eest4.*
6: gb_eest5.*
7: gb_eest6.*
8: gb_gsest1.*
9: gb_gsest2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	777.6	34.9	963	5	B0879880
2	756.2	34.4	881	5	CD518675
3	754.2	33.8	1699	3	AK006549
4	681.8	30.6	711	5	BK095540
5	657	29.5	676	5	BU741926
6	651.6	29.2	659	5	BU730894
7	648.8	29.1	674	5	BU731809
8	641.8	28.8	756	7	CO396179
9	640.4	28.7	666	4	BT480798
10	638	28.6	639	6	CA389968
11	631.2	28.3	636	6	CA397981
12	626.6	28.1	651	5	BU731149
13	582.4	26.1	792	4	BM663028
14	577	25.9	772	4	BM610951
15	575	25.8	585	4	BM707948
16	565.2	25.4	593	2	BE385296
17	528.8	23.7	735	4	BT756228
18	525.8	23.6	537	4	BM691456
19	522	23.4	522	4	BM707649
20	508	22.8	526	4	BM685501
21	495.4	22.2	497	6	CA391418
22	492.4	22.1	930	5	BQ436824
23	492.2	22.1	545	6	CA395098
24	491	22.0	491	7	CV024947

25	485.2	21.8	566	4	BM718338
26	476.4	21.4	582	5	BP379588
27	475.4	21.3	503	4	BM685122
28	475.4	21.3	508	5	BM932117
29	470.8	21.1	474	1	A1190190
30	469.6	21.1	492	5	BU726009
31	467	21.0	484	5	BU741783
32	464.8	20.9	469	1	AA307119
33	449.2	20.2	455	4	BG013943
34	444.6	19.9	601	4	BG951790
35	443	19.9	481	4	BM662514
36	437.4	19.6	531	4	BM685396
37	427.4	19.2	484	7	N31453
38	422	18.9	448	5	BU733195
39	405.4	18.2	419	4	BM718146
40	402.6	18.1	792	7	CK470840
41	401.4	18.0	508	2	BE235901
42	391	17.5	766	2	BE275846
43	388	17.4	713	6	BY714331
44	385.2	17.3	862	7	CK469826
45	379.8	17.0	1067	4	BM562042

ALIGNMENTS

RESULT 1
B0879880
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

B0879880 963 bp mRNA linear EST 16-AUG-2002
AGENCOURT_B241531_lupsk1_dorsal_root_ganglion_Homo_sapiens_CDNA
clone IMAGE:6180559 5', mRNA sequence.
B0879880
B0879880.1 GI:22271888
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 963)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LNLN1363 row: 1 column: 08
High quality sequence stop: 623.
Location/Qualifiers
1..963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_id="lupsk1_dorsal_root_ganglion"
/note="Vector: PCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCGACGCGCG-3' and
5'-GACGAGTCCTGATCGGACGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life

QY	1050	CAGGTCCTCCGTTGGCTGTGGATGAGATGACCAAGACCTGGCTCGGATGAGCCGGAC	1109
Db	481	CAGGTGCTCCCTGTGGCTGTGGATGAGATGACCAAGACCTGGCTCGGATGAGCCGGAC	540
QY	1110	ATGTACTGGAATTAAGCCCGAGCCCAAGCCCCCTTACAGAGCTGCTTCCGCCAAGTTCGGT	1163
Db	541	ATGTACTGGAATTAAGCCCGAGCCCAAGCCCCCTTACAGAGCTGCTTCCGCCAAGTTCGGT	600
QY	1170	CGAGCTCTCTTTATGSGCTCCACCTTCAACATCAGCTCGTGAACAAAGAGAGATGAGTTTC	1223
Db	601	CGAGCTCTCTTTATGSGCTCCACCTTCAACATCAGCTCGTGAACAAAGAGAGATGAGTTTC	660
QY	1230	CAGCCCAATCATGAGAGAGAGAGAGATGCTACCGCTGGCATCATTTGGCCGCTTCTTAGGC	1289
Db	661	CAGCCCAATCATGAGAGAGAGAGAGATGCTACCGCTGGCATCATTTGGCCGCTTCTTAGGC	720
QY	1290	CTGCAGTCCCATGATCATCATCTCTCCCAAGGG--CAAACTCAAGAGCAAACTAGTGGC	1347
Db	721	CTGCAGTCCCATGATCATCATCTCTCCCAAGGGCAAACTCAAGAGCAAACTAGTGGC	780
QY	1348	CCAAAGAGGAA--TCCCTTCTCCACGAGGGCTGCCCCAAAACCAAGGACGCAAA	1404
Db	781	CCCAAGAGGAAATTCCTTCTTCCACGAGGGGCTGCCCAAAAACCCCAAGGACGCC	840
QY	1405	AGAAAGTTAGGGGCCAGAGAACAA	1429
Db	841	AAACCGAAACGTTTATGGGGGCCCA	865
RESULT 3			
AK006549			
LOCUS			
DEFINITION	AK006549	1699 bp	mRNA
			linear HTC 03-APR-2004
			Mus musculus adult male testis cDNA, RIKEN full-length enriched
			library, clone:1700030H21 product:vtell1iform macular dystrophy 2
			homolog (human), full insert sequence.
ACCESSION	AK006549		
VERSION	AK006549.1	GI:12839710	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	1	Carninci, P. and Hayashizaki, Y.	
		High-efficiency full-length cDNA cloning	
JOURNAL		Mech. Enzymol. 303, 19-44 (1999)	
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE			
AUTHORS	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,	
		Ichih, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.	
		Normalization and subtraction of cap-trapper-selected cDNAs to	
		prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL		Genome Res. 10 (10), 1617-1630 (2000)	
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE			
AUTHORS	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,	
		Konno, H., Akiyama, U., Nishi, K., Kikunaka, T., Tashiro, H., Itoh, M.,	
		Sun, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,	
		Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kaishiyagi, K.,	
		Fujiwaka, S., Inoue, K., Togawa, Y., Iwama, M., Ohara, E., Matsuhiki, M.,	
		Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,	
		Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
		RIKEN Integrated sequence analysis (RISA) system-384-format	
		sequencing pipeline with 384 multiplexed sequencer	
JOURNAL		Genome Res. 10 (11), 1757-1771 (2000)	
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE			
AUTHORS	4	The RIKEN Genome Exploration Research Group Phase II Team and the	
		FANTOM Consortium.	

TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE AUTHORS

TITLE The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
REFERENCE AUTHORS Nature 420, 563-573 (2002)

TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, Riken Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.c.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGCGAGACATCCAGAAGCTTTTTCCTTTTTTTTNN 3'], cDNA was prepared by using triethanol thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGAGCAGAGGGCGGCAGATTAACTTCGAGTATTAAATTATCCCCCCCC 3']. cDNA was cleaved with XhoI and SclI. Cloning sites, 5' end: XhoI; 3' end: SclI. Host: SOLR.

FEATURES
SOURCE Location/Qualifiers
 1..1699
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /DB_xref="FANTOM,DB:1700030H21"
 /db_xref="taxon:10090"
 /cclone="1700030H21"
 /sex="male"
 /tissue_type="testis"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="adult"
 1..1699
 /note="Vitellogenin macular dyetrophy 2 homolog (human)"
 (MGI|MGI:1346332)"

ORIGIN

Query Match 33.8%; Score 754.2; DB 3; Length 1699;
 Best Local Similarity 77.3%; Pred. No. 1.6e-179;
 Matches 983; Conservative 0; Mismatches 248; Indels 40; Gaps 4;

```

OY      273 GAACAACGCTGATGTTTGAAAACGTACCTGTGATTCGACACAGCTACCTACCTATC 332
Db       1 GATCAGACAGCTGTTGTTGAAGAGCTGCTCTGACTGCAACAGCTACCATTCAGCTATC 60
OY      333 CCATTTCTTCGTGCTGAGGCTTCTACAGTACCGTGCTGTGACCCTGGTGGAACG 392
Db       61 CCTAATCTCTTCGTCTGGGTTTCTATGATTACATGATGGTAGGCCGCTGGTAGGACG 120
OY      393 TAGGAAACCTGCCTGGGCCGACCGCTCATGATGAGCTGTGTGTGGGCTTCGTGAAGGC 452

```


QY 1452 GTGAGCGCTTCAAGTCTGGCCCACTGATCAGAGGCGAGGCTACTACAGTGGCCCAAG 1511
 |||||
 DB 121 GTGAGCGCTTCAAGTCTGGCCCACTGATCAGAGGCGAGGCTACTACAGTGGCCCAAG 180
 |||||
 QY 1512 ACGCCCTCAGCGCCCACTCCCATGTTCTTCCCTAGAACCATGAGCGCGCTCAAGCTT 1571
 |||||
 DB 181 ACGCCCTCAGCGCCCACTCCCATGTTCTTCCCTAGAACCATGAGCGCGCTCAAGCTT 240
 |||||
 QY 1572 CACAGTGTACAGGATAGACACCAAGACAAAGCTTAAAGACTGTGATGTTCTGGGGCC 1631
 |||||
 DB 241 CACAGTGTACAGGATAGACACCAAGACAAAGCTTAAAGACTGTGATGTTCTGGGGCC 300
 |||||
 QY 1632 AAGAAAAGTTTGAATGCTCTCAGAGAGGATGGGGCTTGATGAGACACCCAGAAAGTA 1691
 |||||
 DB 301 AAGAAAAGTTTGAATGCTCTCAGAGAGGATGGGGCTTGATGAGACACCCAGAAAGTA 360
 |||||
 QY 1692 TCTCAAGTGAAGAGAAACCTGTGAGTTTAACTGACGATATGCCAGAGATCCCGAA 1751
 |||||
 DB 361 TCTCAAGTGAAGAGAAACCTGTGAGTTTAACTGACGATATGCCAGAGATCCCGAA 420
 |||||
 QY 1752 AATCACTCAAGAAACCTTTGGAAACATCACAACCAACATACACACTCACTCAAGAT 1811
 |||||
 DB 421 AATCACTCAAGAAACCTTTGGAAACATCACAACCAACATACACACTCACTCAAGAT 480
 |||||
 QY 1812 CACATGATCCTTATTTGGGCGCTTGAAAACAGGGATGAGACATTTCTTACCTGCTTCC 1871
 |||||
 DB 481 CACATGATCCTTATTTGGGCGCTTGAAAACAGGGATGAGACATTTCTTACCTGCTTCC 540
 |||||
 QY 1872 TAATGGGAGATGCTTCGCGACGAGGCTCTCACTGTGTGTACACGAGGAGCACTGATC 1931
 |||||
 DB 541 TAATGGGAGATGCTTCGCGACGAGGCTCTCACTGTGTGTACACGAGGAGCACTGATC 600
 |||||
 QY 1932 CAGTCAACGACCATACAGTGTTCACACTGAAAGACGTGTCTTCAACAACGCTGAATCAA 1991
 |||||
 DB 601 CAGTCAACGACCATACAGTGTTCACACTGAAAGACGTGTCTTCAACAACGCTGAATCAA 660
 |||||
 QY 1992 TGGTACCTTAATGATTAATAATCCCGACTACTTCA 2028
 |||||
 DB 661 TGGTACCTTAATGATTAATAATCCCGACTACTTCA 697
 |||||

RESULT 5
 LOCUS BU741926/c 676 bp mRNA linear EST 10-OCT-2002
 DEFINITION UI-E-EO1-aj-a-j-23-0-UI-s1 UI-E-EO1 Homo sapiens cDNA clone
 UI-E-EO1-aj-a-j-23-0-UI 3', mRNA sequence.
 ACCESSION BU741926
 VERSION BU741926.1 GI:23687736
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 676)
 Bonaudo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL MEDLINE
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research

FEATURES
 SOURCE
 Location/Qualifiers
 1..676
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-EO1-aj-a-j-23-0-UI"
 /tissue_type="fetal eye"
 /dev_stage="fetal"
 /clone_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-EO1"
 /note="Organ: eye; Vector: pTTT3-Pac (Pharmacia) with a
 modified polylinker; Site 1: Scorer I; Site 2: Not I;
 UI-E-EO1 is a normalized cDNA library containing the
 following tissue(s): fetal eye. The library was
 constructed according to Bonaudo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pTTT3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. This sequence tag for this library is
 CGCGTATACG. This library was created for the program, Gene
 Discovery in the Visual System, supported by National Eye
 Institute (NEI).
 TAG_TISSUE=human fetal eye
 TAG_LIB=UI-E-EO1
 TAG_SEQ=CCGCTATACC"

ORIGIN

Query Match 29.5%; Score 657; DB 5; Length 676;
 Best Local Similarity 99.1%; Pred. No. 5.6e-155;
 Matches 671; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
 QY 1543 CCTTGAACCATTCAGCGCCGCTCAAGGCTTCACTGTCACAGGATGACCAAGCA 1602
 |||||
 DB 676 CCTTGAACCATTCAGCGCCGCTCAAGGCTTCACTGTCACAGGCTGACCAAGCA 618
 |||||
 QY 1603 AAAGCTTAAAGACGTGAGTCTGGGCGCAAGAAAGTTTGAATGCTCTCAGAGACG 1662
 |||||
 DB 617 AAAGCTTAAAGACGTGAGTCTGGGCGCAAGAAAGTTTGAATGCTCTCAGAGACG 558
 |||||
 QY 1663 ATGGGGCCTTGATGAGACACCCAGAGATATCTCAAGTGAAGAGAAACCTGTGAAGTTTA 1722
 |||||
 DB 557 ATGGGGCCTTGATGAGACACCCAGAGATATCTCAAGTGAAGAGAAACCTGTGAAGTTTA 498
 |||||
 QY 1723 ACCTGAAGGATATGCCAGAGATCCCCGAAATATCATCTTGAAGAACTTTGGAACAATAC 1782
 |||||
 DB 497 ACCTGAAGGATATGCCAGAGATCCCCGAAATATCATCTTGAAGAACTTTGGAACAATAC 438
 |||||
 QY 1783 CAACCAACATACACATACACTCAAGATCATGATCCTTATTTGGGCTTTGAAAAACA 1842
 |||||
 DB 437 CAACCAACATACACATACACTCAAGATCATGATCCTTATTTGGGCTTTGAAAAACA 378
 |||||
 QY 1843 GGGATGAAGACATTTCTTAACCTGCTTCTTAATGGGAGTCTTGCCAGCCAGTCTCA 1902
 |||||
 DB 377 GGGATGAAGACATTTCTTAACCTGCTTCTTAATGGGAGTCTTGCCAGCCAGTCTCA 318
 |||||
 QY 1903 CCTGTGTATACACAGACAGGACTGATCCAGTCCAGCCATACAGTGTCCACATGAA 1962
 |||||
 DB 317 CCTGTGTATACACAGACAGGACTGATCCAGTCCAGCCATACAGTGTCCACATGAA 258
 |||||
 QY 1963 GAACGTGTCTCAACAAGCCTGAATCAATGAGTTAGTTAATAGATTAATATCCAGACT 2022
 |||||
 DB 257 GAACGTGTCTCAACAAGCCTGAATCAATGAGTTAGTTAATAGATTAATATCCAGACT 198
 |||||
 QY 2023 ACTTCAAGCTTTAATGCTTTTATTCATTAATAAACTGTGAAGTGAACCATTTGA 2082
 |||||
 DB 197 ACTTCAAGCTTTAATGCTTTTATTCATTAATAAACTGTGAAGTGAACCATTTGA 138
 |||||

QY 2083 AACATTAACTGAGACTGTGATTCAGAGTCCGGAACCCCTTAGTTCTATCTGAATCCAG 2142
 Db 137 AACATTAACTGAGACTGTGATTCAGAGTCCGGAACCCCTTAGTTCTATCTGAATCCAG 78
 QY 2143 ACAGCCACACCTTAGTACTGCTCCCAACTAATGATTAATTAATCAATACTCGTTA 2202
 Db 77 ACAGCCACACCTTAGTACTGCTCCCAACTAATGATTAATTAATCAATACTCGTTA 18
 QY 2203 AAAAAAAAAAAAAAAAAA 2219
 Db 17 AAAAAAAAAAAAAAAAAA 1

RESULT 6
 LOCUS BU730894/c 659 bp mRNA linear EST 09-OCT-2002
 DEFINITION UI-E-C11-afw-i-11-0-UI s1 UI-E-C11 Homo sapiens cDNA clone
 ACCESSION BU730894
 VERSION BU730894.1 GI:23655244
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 659)
 Bernaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 FORWARD
 PolyA=yes

FEATURES
 source location/Qualifiers
 1..659
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-afw-i-11-0-UI"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_id="UI-E-C11"
 /note="Organ: eye; Vector: pRTT3-Pac (pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-C11 is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Bernaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pRTT3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 (dt)18 tail. The sequence tag for this library is ACCTA.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute

ORIGIN
 Query Match 29.2%; Score 651.6; DB 5; Length 659;
 Best Local Similarity 99.2%; Pred. No. 1.3e-153;
 Matches 654; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 TAG_SEQ=ACCTA"

(NEI).
 TAG_TISSUE=RPE and Choroid
 TAG_LIB=UI-E-C11
 TAG_SEQ=ACCTA"

QY 1561 CGTCAAGCTTCAAGTGTCAAGAGTACAGCAACCAAGCAAGCTTAAGACTGTGA 1620
 Db 659 CGTCAAGCTTCAAGTGTCAAGAGTACAGCAACCAAGCAAGCTTAAGACTGTGA 600
 QY 1621 GTTCTGGGGCCCAAGAAAGTTTGAATTCCTCTCAAGAGCGATGGGCTTGATGAGAC 1680
 Db 599 GTTCTGGGGCCCAAGAAAGTTTGAATTCCTCTCAAGAGCGATGGGCTTGATGAGAC 540
 QY 1681 ACCCAAGATATCTCAAGTGAAGGAGAAACCTGTGAGTTTAACTGAGGATATGCCAG 1740
 Db 539 ACCCAAGATATCTCAAGTGAAGGAGAAACCTGTGAGTTTAACTGAGGATATGCCAG 480
 QY 1741 AGATCCCGGAAATCACTTCAAGAAACCTTTGGAACATACCAACCAATACACATA 1800
 Db 479 AGATCCCGGAAATCACTTCAAGAAACCTTTGGAACATACCAACCAATACACATA 420
 QY 1801 CACTCAAGATCAATGATATCTTATTTGGGCTTGGAAACAGGAGATGAAGACATTCTT 1860
 Db 419 CACTCAAGATCAATGATATCTTATTTGGGCTTGGAAACAGGAGATGAAGACATTCTT 360
 QY 1861 AACCTGCTTCAATGAGGATATCTTGGCCAGGCTCTCACTGTGTGACACACAGA 1920
 Db 359 AACCTGCTTCAATGAGGATATCTTGGCCAGGCTCTCACTGTGTGACACACAGA 300
 QY 1921 GACACTGATTCAGTACAGCATACAGCTGCCACTGGAAGAACGCTCTACACAG 1980
 Db 299 GACACTGATTCAGTACAGCATACAGCTGCCACTGGAAGAACGCTCTACACAG 240
 QY 1981 CCTGAATCAATGTTAGTTATATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
 Db 239 CCTGAATCAATGTTAGTTATATGATTAATTAATTAATTAATTAATTAATTAATTAAT 180
 QY 2041 TTTTATTCATTAATAAATCTGGAAGCTGACGAACCATTTGGAACATTTAATCTGACACTC 2100
 Db 179 TTTTATTCATTAATAAATCTGGAAGCTGACGAACCATTTGGAACATTTAATCTGACACTC 120
 QY 2101 TGGATTTCAGATCGGGAACCTTATGTTCTATCTGAATTCAGACGACACACCTTAGTAT 2160
 Db 119 TGGATTTCAGATCGGGAACCTTATGTTCTATCTGAATTCAGACGACACACCTTAGTAT 60
 QY 2161 ACTGCCCAACTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2219
 Db 59 ACTGCCCAACTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1

RESULT 7
 LOCUS BU731809/c 674 bp mRNA linear EST 09-OCT-2002
 DEFINITION UI-E-C11-afw-i-11-0-UI s1 UI-E-C11 Homo sapiens cDNA clone
 ACCESSION BU731809
 VERSION BU731809.1 GI:23657073
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 674)
 Bernaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477

PUBMED 8889548
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 FORWARD
 PolyA+ies.

FEATURES
 source
 Location/Qualifiers
 1..674
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-af1-g-11-0-UI"
 /issue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-C11"
 /note="Organ: eye; Vector: pT773-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-C11 is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT773-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is ACCTA.
 This library was created for the program, Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI).
 TAG_TISSUE=RPE and Choroid
 TAG_LIB=UI-E-C11
 TAG_SEQ=ACCTA"

ORIGIN
 Query Match 29.1%; Score 648.8; DB 5; Length 674;
 Best Local Similarity 97.9%; Pred. No. 6.7e-153;
 Matches 656; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

1548 GAACCATCAGGCGCGTCAAGCTTCAAGTGTCAAGGATAGACACCAAGACAAAAGC 1607
 670 GCACGAGGAGGCGCGTCAAGCTTCAAGTGTCAAGGATAGACACCAAGACAAAAGC 611
 1608 TTTAAGACTGTGAGTTCGTGGGCGCAAGAAAGTTTGAATGCTCTCAAGAGGATGGG 1667
 610 TTTAAGACTGTGAGTTCGTGGGCGCAAGAAAGTTTGAATGCTCTCAAGAGGATGGG 551
 1668 GCCTTGATGAGCAACCCAGAGTATCTCAAGTGAAGAGAAAGCTGTGAGTTTAACTTG 1727
 550 GCCTTGATGAGCAACCCAGAGTATCTCAAGTGAAGAGAAAGCTGTGAGTTTAACTTG 491
 1728 ACGGATATGCCAGAGATCCCGAAATATCACTCAAGAACTTTGGAACAATCAACAAC 1787
 490 ACGGATATGCCAGAGATCCCGAAATATCACTCAAGAACTTTGGAACAATCAACAAC 431
 1788 AACCTACACATTAACCTCAAGATCAATGATCTTATTTGGGCTTGGAAAAAGGAT 1847
 430 AACCTACACATTAACCTCAAGATCAATGATCTTATTTGGGCTTGGAAAAAGGAT 371
 1848 GAACACATTCCTTAACCTGTTCTTAATGGGATGCTTCGACAGCAGATCCTCAACCTGT 1907

Db 370 GAACACATTCCTTAACCTGCTCCTTAATGGGATGCTTCGACAGCAGATCCTCAACCTGT 311
 Qy 1208 GTGTACACCGGAGAGACATGATCCAGTCAAGCATACAGCTGTCAACTGAAGAAGC 1967
 Db 310 GTGTACACCGGAGAGACATGATCCAGTCAAGCATACAGCTGTCAACTGAAGAAGC 251
 Qy 1968 TGTCTCTCAACAGGCTGAATCAATGAGTTAGCTTAATAGATTAATAATCCAGACTACTTC 2027
 Db 250 TGTCTCTCAACAGGCTGAATCAATGAGTTAGCTTAATAGATTAATAATCCAGACTACTTC 191
 Qy 2028 AGCCTTAAATGCTTTTATTCATTAATAAACTGTGAAGCTAGACTGAACATTGGAACAT 2087
 Db 190 AGCCTTAAATGCTTTTATTCATTAATAAACTGTGAAGCTAGACTGAACATTGGAACAT 131
 Qy 2088 TTAACCTGAGCTCGATGATTCAGAGTCCGGAACCTTGTCTATCTGAATCAAGACAGC 2147
 Db 130 TTAACCTGAGCTCGATGATTCAGAGTCCGGAACCTTGTCTATCTGAATCAAGACAGC 71
 Qy 2148 CACACCTTAGTACTGCTCCCAACTAATGAGTTTAAATTAACAATTAATCTGTTAAAAA 2207
 Db 70 CACACCTTAGTACTGCTCCCAACTAATGAGTTTAAATTAACAATTAATCTGTTAAAAA 11
 Qy 2208 AAAAAAAAAA 2217
 Db 10 AAAAAAAAAA 1

RESULT 8
 CO396179 756 bp mRNA linear EST 01-JUL-2004
 LOCUS AGENCOURT 27872871 NIH MGC 212 Homo sapiens cDNA clone
 DEFINITION IMAGE:30923539 5', mRNA sequence.
 ACCESSION CO396179
 VERSION CO396179.1 GI:49578095
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 756)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabds-rc@mail.nih.gov
 Tissue Procurement: Mary Hendrix
 CDNA Library Preparation: M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM164 row: 1 column: 20
 High quality sequence. Stop: 571.
 Location/Qualifiers
 1..756
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30923539"
 /issue_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH MGC 212"
 /note="Organ: Lung; Vector: pTX-Lac; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated

FEATURES
 source

with Sc0r I adaptor, digested with Not I and then cloned directionally into pUX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GATAGGCCCA. Tissue was provided by Mary Hendrix."

JOURNAL
COMMENT
Unpublished (2001)
Contact: Swaroop, A.
Department of Ophthalmology and Visual Sciences
Kellogg Eye Center, University of Michigan
540 KEC, 1000 Mail St., Ann Arbor, MI 48105, USA
Tel: 734 615 2246
Fax: 734 647 0228
Email: swaroop@umich.edu

Query Match
Best Local Similarity 28.8%; Score 641.8; DB 7; Length 756;
Matches 646; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

PCR Primers
FORWARD: M13/PUC-Reverse - ccagctcagcagctgttaaacg
BACKWARD: M13/PUC-Forward - agcgataacattccacagc
Seq primer: M13/PUC-Reverse.

417 CGCCTCATGAGCTGGGTGGGCTTCGCGCAAGCAGACAGACAGAGCCGCTGCTG 476
8 CGCCTCATGAGCTGGGTGGGCTTCGCGCAAGCAGACAGACAGAGCCGCTGCTG 67
477 CGCGGAGAGCTGATCCGCTACGCCAACCTGGGCAACGCTGCTATCCGCGAGGCTGAC 536
68 CGCGGAGAGCTGATCCGCTACGCCAACCTGGGCAACGCTGCTATCCGCGAGGCTGAC 127
537 ACCGAGTCTACAGAGGCTTCGCCAGCGCCAGACCTGGTCAAGAGGCTTTATGACT 596
128 ACCGAGTCTACAGAGGCTTCGCCAGCGCCAGACCTGGTCAAGAGGCTTTATGACT 187
597 CGCGGAGAGCTGATCCGCTACGCCAACCTGGGCAACGCTGCTATCCGCGAGGCTGAC 656
188 CGCGGAGAGCTGATCCGCTACGCCAACCTGGGCAACGCTGCTATCCGCGAGGCTGAC 247
657 TGGGTGTGGTTCGCAACCTGCTCAATGAGCGCTGGTGGAGTCCGAGACCT 716
248 TGGGTGTGGTTCGCAACCTGCTCAATGAGCGCTGGTGGAGTCCGAGACCT 307
717 ATCTGCTCCAGAGCTGCTGTAAGAGATGATGACACCTGGGTAAGAGTGGAGACCTG 776
308 ATCTGCTCCAGAGCTGCTGTAAGAGATGATGACACCTGGGTAAGAGTGGAGACCTG 367
777 TATGCTACAGCTGATGATATCCACCTGGTGTATACAGAGTGGTGAATGTCGCGTG 836
368 TATGCTACAGCTGATGATATCCACCTGGTGTATACAGAGTGGTGAATGTCGCGTG 427
837 TATGCTACAGCTGATGATATCCACCTGGTGTATACAGAGTGGTGAATGTCGCGTG 896
428 TATGCTACAGCTGATGATATCCACCTGGTGTATACAGAGTGGTGAATGTCGCGTG 487
897 CCTGAGCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 956
488 CCTGAGCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 547
957 GTTGGCTGCTGTAAGAGTGGCAGAGAGCTCATCAACCTTTGAGAGAGATGATGAT 1016
548 GTTGGCTGCTGTAAGAGTGGCAGAGAGCTCATCAACCTTTGAGAGAGATGATGAT 607
1017 TTTGAGACCACTGATGTCAGAGAGATTTGAGAGTGTCCCTGTTGGCTGTG 1070
608 TTTGAGACCACTGATGTCAGAGAGATTTGAGAGTGTCCCTGTTGGCTGTG 661

FEATURES
source
1..666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Native Retinal Pigment Epithelium sheets"
/dev_stage="juvenile"
/clone_lib="Human Retinal Pigment Epithelium (2)"
/note="Organ: Retina; Vector: pSport1"

ORIGIN

Query Match
Best Local Similarity 28.7%; Score 640.4; DB 4; Length 666;
Matches 652; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

10 CCACAGCCCTAGTCGCGCAGACCTTCGCGGATCATTCGACCCACCTGGAACCCACCTG 69
12 CCACAGCCCTAGTCGCGCAGACCTTCGCGGATCATTCGACCCACCTGGAACCCACCTG 71
70 ACCGAGCCCACTGCTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 129
72 ACCGAGCCCACTGCTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 131
130 TGGCTAATGCGCGCTTAAAGCTCTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 189
132 TGGCTAATGCGCGCTTAAAGCTCTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 191
190 ACAGCTGCTAATGCGCGCTTAAAGCTCTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 249
192 ACAGCTGCTAATGCGCGCTTAAAGCTCTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 251
250 TTTAAGGCTGCGCGCTTAAAGCTCTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309
252 TTTAAGGCTGCGCGCTTAAAGCTCTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 311
310 GCGACAGCTAATGCGCGCTTAAAGCTCTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 369
312 GCGACAGCTAATGCGCGCTTAAAGCTCTTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 371
370 TCGTGAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429
372 TCGTGAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 431
430 TGGTGTGGGCTTCTGCGAGGAG 489
432 TGGTGTGGGCTTCTGCGAGGAG 491
490 TCGCTAAGCGCAACCTGGGCAACGCTGCTATCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 549
492 TCGCTAAGCGCAACCTGGGCAACGCTGCTATCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 551
550 AGGCGTTCGCCAGCGCCAG 609
552 AGGCGTTCGCCAGCGCCAG 610
610 ACCAGTTGAGAGAACTGAGCTTACCAACCAATGTTCTGGGTGCTGCTGCTGCTGCTGCTG 665
611 ACCAGTTGAGAGAACTGAGCTTACCAACCAATGTTCTGGGTGCTGCTGCTGCTGCTGCTG 666

RESULT 9
B1480798
LOCUS B1480798 666 bp mRNA linear EST 28-FEB-2002
DEFINITION H2RPE-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA
sequence.
B1480798
ACCESSION B1480798.1 GI:18998607
VERSION B1480798.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 666)
Yuan, Y., MacNee, S.P., Hughes, B. and Swaroop, A.
Towards an expression profile of native human retinal pigment
epithelium: Identification of a non-redundant set of more than 1100

/clone lib="Human Retinal pigment epithelium/choroid cDNA (un-normalized, unamplified): cs"
 /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor eyes (75-80 years old) yielded approximately 600 mg of dissected RPE/choroid tissue. This in turn yielded 340 ug of total RNA and 7 ug of mRNA. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies (Rockville, MD; now part of Invitrogen Corp), essentially following the protocols of the Superscript Plasmid System (Invitrogen Corp).
 <http://www.invitrogen.com/>. The library code designation was cs. For this library, cDNA inserts were cloned into the NotI/MluI sites of the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 28.3%; Score 631.2; DB 6; Length 636;
 Best Local Similarity 99.5%; Pred. No. 1.9e-148;
 Matches 633; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1125 CCCGAGCCACAGCCCCCTACACAGCTGCTCCGCCCGTCCGTCAGCCCTTTATG 1184
 DB 1 CCCGAGCCACAGCCCCCTACACAGCTGCTCCGCCCGTCCGTCAGCCCTTTATG 60
 QY 1185 GCGTCCACCTTCAACATCAGCTGGAACAAAGAGATGAGTTCCAGCCCATCAGAG 1244
 DB 61 GCGTCCACCTTCAACATCAGCTGGAACAAAGAGATGAGTTCCAGCCCATCAGAG 120
 QY 1245 GACGAGGAGATGCTCAGGCTGGCATCATGTCGCCCTTCTTACGCTTCCAGTCCATGAT 1304
 DB 121 GACGAGGAGATGCTCAGGCTGGCATCATGTCGCCCTTCTTACGCTTCCAGTCCATGAT 180
 QY 1305 CACCATCTCTCCAGGGGCAAACTCAAGAGACCAAACTACTGTGGCCCAAGAGGAATCCCTT 1364
 DB 181 CACCATCTCTCCAGGGGCAAACTCAAGAGACCAAACTACTGTGGCCCAAGAGGAATCCCTT 240
 QY 1365 CTCGACGAGGGCTGCGCCCAAAACCAAGGACGACCAAGCAAGCTTAAAGGGGCGCAGGAA 1424
 DB 241 CTCGACGAGGGCTGCGCCCAAAACCAAGGACGACCAAGCAAGCTTAAAGGGGCGCAGGAA 300
 QY 1425 GACAAACAAGGCTGAGAGCTTAAAGCTGTGAGCGCTTCAAGTCTGGCCCACTGTATCAG 1484
 DB 301 GACAAACAAGGCTGAGAGCTTAAAGCTGTGAGCGCTTCAAGTCTGGCCCACTGTATCAG 360
 QY 1485 AGGCGAGGCTACTACAGTGTCCCAAGAGCGCTTCAAGCGCCCACTCCCATGTTCTTCCC 1544
 DB 361 AGGCGAGGCTACTACAGTGTCCCAAGAGCGCTTCAAGCGCCCACTCCCATGTTCTTCCC 420
 QY 1545 CTGAACCATCAGGCGCGTCAAGAGCTTCAAGTGTCAAGGCAATGACACCAAGACAA 1604
 DB 421 CTGAACCATCAGGCGCGTCAAGAGCTTCAAGTGTCAAGGCAATGACACCAAGACAA 480
 QY 1605 AGCTTAAAGACTGTAGTTCGTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGGAT 1664
 DB 481 AGCTTAAAGACTGTAGTTCGTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGGAT 540
 QY 1665 GGGGCTTGTGAGGACCAAGAGATCTCAAGTGAAGGAGAAAGCTGTGAAGTTAAC 1724
 DB 541 GGGGCTTGTGAGGACCAAGAGATCTCAAGTGAAGGAGAAAGCTGTGAAGTTAAC 600
 QY 1725 CTGACGATATGCGACAGATCCCGGAAATCACCCTC 1760
 DB 601 CTGACGATATGCGACAGATCCCGGAAATCACCCTC 636

RESULT 12
 BU731149/c 651 bp mRNA linear EST 09-OCT-2002
 LOCUS BU731149
 DEFINITION UI-E-CII-afq-j-20-0-UI.61 UI-E-CII Homo sapiens cDNA clone
 ACCESSION BU731149
 VERSION BU731149.1 GI:23655752

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 651)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

Location/Qualifiers
 1..651
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-CII-afq-j-20-0-UI"
 /issue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-CII"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-CII is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACTTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
 TAG_TISSUE=RPE and Choroid
 TAG_LIB=UI-E-CII
 TAG_SEQ=ACCTA"

ORIGIN

Query Match 28.1%; Score 626.6; DB 5; Length 651;
 Best Local Similarity 99.2%; Pred. No. 2.8e-147;
 Matches 629; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1390 ACAAGGACGCAACAGACCTTAGGGCCAGAGACCAACAAGGCTGGAAGCTTAAG 1449
 DB 651 ACAAGGACGCAACAGACCTTAGGGCCAGAGACCAACAAGGCTGGAAGCTTAAG 592
 QY 1450 CTGTGAGCGCTTCAAGCTGTGAGCCCACTGTATCAGAGGCCAGGCTACTAAGTCCAC 1509
 DB 591 CTGTGAGCGCTTCAAGCTGTGAGCCCACTGTATCAGAGGCCAGGCTACTAAGTCCAC 532
 QY 1510 AGAGCGCCCTGAGCGCCCACTCCCATGTTCTTCCCTTCAAGACCATCAAGCGCGCAAGC 1569
 DB 531 AGAGCGCCCTGAGCGCCCACTCCCATGTTCTTCCCTTCAAGACCATCAAGCGCGCAAGC 472

DEFINITION 601303662F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3638175 5',
 mRNA sequence.
 BE410951
 VERSION BE410951.1 GI:9347401
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 773)
 NIH-MGC http://mgc.ncl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINTL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINTL at: image.lnl.gov
 Plate: L1CM38 row: 1 column: 16
 High quality sequence stop: 662.

FEATURES
 source
 1..773
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3638175"
 /tissue_type="choriocarcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Placenta; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 25.9%; Score 577; DB 2; Length 773;
 Best Local Similarity 99.0%; Pred. No. 1.1e-134;
 Matches 591; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
 1253 GGATGCTACCGCTGCGATCATTTGGCCGCTTCTTGAAGCTGATGATCCATCC 1312
 1 GGATGCTACCGCTGCGATCATTTGGCCGCTTCTTGAAGCTGATGATCCATCC 60
 1313 TCCCAAGGCAAACTCAAGACCAAACTAAGTGGCCCAAGAGGGAATCCCTTCCACCA 1372
 61 TCCCAAGGCAAACTCAAGACCAAACTAAGTGGCCCAAGAGGGAATCCCTTCCACCA 120
 1373 GGGCCCTGCC-AAAAACCAAGGACCAAGCAAAAGTTAGGGCCAGGAAGACACA 1431
 121 GGGCCCTGCCAAAAACCAAGGACCAAGCAAAAGTTAGGGCCAGGAAGACACA 180
 1432 AGGCTTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTGTGGCCCACTGTATCAGAGCCG 1491
 181 AGGCTTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTGTGGCCCACTGTATCAGAGCCG 240
 1492 GCTACTACATGCCCCCAAGCGCCCTGAGCCCACTCCCATGTTCTTCCCTAGAAC 1551
 241 GCTACTACATGCCCCCAAGCGCCCTGAGCCCACTCCCATGTTCTTCCCTAGAAC 300
 1552 CATCAGCGCGCTCAAAAGCTTCAAGTGTACAGGGCAATACCAAAAGCAAAAGCTTAA 1611
 301 CATCAGCGCGCTCAAAAGCTTCAAGTGTACAGGGCAATACCAAAAGCAAAAGCTTAA 360
 1612 AGACTGTGAGTTCTGGGGCCCAAGAAAGTTTGAATGTGCTCAGAGAGCGATGGGGCT 1671
 361 AGACTGTGAGTTCTGGGGCCCAAGAAAGTTTGAATGTGCTCAGAGAGCGATGGGGCT 420

QY 1672 TGATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAAACTGTGAGTTAACTGACGG 1731
 DB 421 TGATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAAACTGTGAGTTAACTGACGG 480
 QY 1732 ATATGCCAGAGATCCCGGAAATATCACTCAAAAGAACTTTGGAACATCCACCAACA 1791
 DB 481 ATATGCCAGAGATCCCGGAAATATCACTCAAAAGAACTTTGGAACATCCACCAACA 540
 QY 1792 TACACACTACACTCAAAAGATCAATGATGATCTTATTTGGGCTTGAAGAAACAGGGATG 1848
 DB 541 TACACACTACACTCAAAAGATCAATGATGATCTTATTTGGGCTTGAAGAAACAGGTCTG 597

RESULT 15
 BM707948
 LOCUS
 DEFINITION BM707948 585 bp mRNA linear EST 28-FEB-2002
 UI-E-C11-afg-9-11-0-UI.r1 UI-E-C11 Homo sapiens cDNA clone
 BM707948
 VERSION BM707948.1 GI:19021206
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 585)
 Bernaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES

source
 1..585
 Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-afg-9-11-0-UI"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: eye; Vector: pUT73-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-E-C11 is a normalized cDNA library containing the
 following tissue(s): RPE and Choroid. The library was
 constructed according to Bernaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. The library was
 synthesized with a primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pUT73-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is ACCTA.
 This library was created for the program Gene Discovery
 in the Visual System, supported by National Eye Institute
 (NEI)."

ORIGIN

Query Match 25 8%; Score 575; DB 4; Length 585;
Best Local Similarity 98.8%; Pred. No. 3.2e-134;
Matches 578; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 1556 AGCCCGCTCAAGCTTCACAGCATGACACCAAGACAAAAGCTTAAAGAC 1615
    |||
Db 1 AGCCCGCTCAAGCTTCACAGCATGACACCAAGACAAAAGCTTAAAGAC 60
    |||

QY 1616 TGTAGTTCTGGGGCCCAAGAAAGTTGAAATGCTCTCAGAGAGCGATGGGCTTGAT 1675
    |||
Db 61 TGTAGTTCTGGGGCCCAAGAAAGTTGAAATGCTCTCAGAGAGCGATGGGCTTGAT 120
    |||

QY 1676 GGAGACCCCAAGAAATGTTCAAGTGAAGAGAAAAGTGAAGTTTAACTGACGATAT 1735
    |||
Db 121 GGAGACCCCAAGAAATGTTCAAGTGAAGAGAAAAGTGAAGTTTAACTGACGATAT 180
    |||

QY 1736 GCCAGAGATCCCGGAAATCACTCAAGAAACCTTTGGACAAATCACCAACCAATPACA 1795
    |||
Db 181 GCCAGAGATCCCGGAAATCACTCAAGAAACCTTTGGACAAATCACCAACCAATPACA 240
    |||

QY 1796 CACTACACTCAAGATCATGATCCTTATTGGGCTTTGAAAAACAGGATGAAGACA 1855
    |||
Db 241 CACTACACTCAAGATCATGATCCTTATTGGGCTTTGAAAAACAGGATGAAGACA 300
    |||

QY 1856 TTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGTCTTCACTGTGTACAC 1915
    |||
Db 301 TTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGTCTTCACTGTGTACAC 360
    |||

QY 1916 CAGCAGGCACTGATCCAGTCAACAGCCATACAGGCTGTCACACTGAAGAACGTGCTTAC 1975
    |||
Db 361 CAGCAGGCACTGATCCAGTCAACAGCCATACAGGCTGTCACACTGAAGAACGTGCTTAC 420
    |||

QY 1976 AACAGCCTGAATCAATGTTAGCTTAATAGATAAAATCCAGACTACTTCAGCCTTTA 2035
    |||
Db 421 AACAGCCTGAATCAATGTTAGCTTAATAGATAAAATCCAGACTACTTCAGCCTTTA 480
    |||

QY 2036 ATGCTTTTATTCATTAATAAACTGTGAAGCTAGACTGAACATTTGAAAACATTTAACTCA 2095
    |||
Db 481 ATGCTTTTATTCATTAATAAACTGTGAAGCTAGACTGAACATTTGAAAACATTTAACTCA 540
    |||

QY 2096 GACTCTGATTCAGAGTCGGGACCCCTTAAGTTCTATCTGAATCCA 2140
    |||
Db 541 GACTCTGATTCAGAGTCGGGACCCCTTAAGTTCTATCTGAATCCA 585
    |||
```

Search completed: March 26, 2005, 13:32:03
Job time : 7175.13 secs

This Page Blank (uspto)

CC mutated is responsible for Best's macular dystrophy (BMD).
CC Polymorphisms encoding GAGE are useful for diagnosing whether a
CC patient carries a mutation in the GAGE gene. Normal and mutated GAGE
CC proteins are useful for identifying activators and/or inhibitors of these
CC proteins, in order to treat BMD. The GAGE gene offers a simpler and
CC cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
CC age-related macular dystrophy
XX

Sequence 2229 BP; 575 A; 646 C; 532 G; 476 T; 0 U; 0 Other;

Query Match 100.0%; Score 2229; DB 2; Length 2229;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CAGGAGTCCACAGCTATGTCGACAGCTTCTGGAGATCATGAGCCACCTGGA 60
Db 1 CAGGAGTCCACAGCTATGTCGACAGCTTCTGGAGATCATGAGCCACCTGGA 60
QY 61 CCCACCTGACCCCAAGCCCACTGTCAGAGCCCACTGGCCATGACCATCATTTACA 120
Db 61 CCCACCTGACCCCAAGCCCACTGTCAGAGCCCACTGGCCATGACCATCATTTACA 120
QY 121 CAGGCAAGTGGTATGCGCCGCTTGGCTCTTCCCGCTGCTGTGTGGCGGG 180
Db 121 CAGGCAAGTGGTATGCGCCGCTTGGCTCTTCCCGCTGCTGTGTGGCGGG 180
QY 181 GGAGCATCTACAAGCTGCTATATGCGAGGTTCTTATCTTCCGCTGCTGCTATCA 240
Db 181 GGAGCATCTACAAGCTGCTATATGCGAGGTTCTTATCTTCCGCTGCTGCTATCA 240
QY 241 TCCGCTTTATTTATAGGCTGGCCCTTACAGGAACAAACAGCTGATGTTTGAAGAACTGA 300
Db 241 TCCGCTTTATTTATAGGCTGGCCCTTACAGGAACAAACAGCTGATGTTTGAAGAACTGA 300
QY 301 CTCTGATATGAGCAAGCTATCATCCAGCTCATCCCATTTCTTCTGCTGGCTTTTACG 360
Db 301 CTCTGATATGAGCAAGCTATCATCCAGCTCATCCCATTTCTTCTGCTGGCTTTTACG 360
QY 361 TGAGCTGTGTCGTGACCCGCTGTGTGAACAGATACAGAACCTGCGTGGCCGACGCGC 420
Db 361 TGAGCTGTGTCGTGACCCGCTGTGTGAACAGATACAGAACCTGCGTGGCCGACGCGC 420
QY 421 TCATAGAGCTGTGTCGTGCTTCTGTCGAAGGCAAGCAAGGCGGCTGTGCGGC 480
Db 421 TCATAGAGCTGTGTCGTGCTTCTGTCGAAGGCAAGCAAGGCGGCTGTGCGGC 480
QY 481 GCAGCTCATCGCTACAGCCCAACCTGGGCAAGTCTCATCTGCGAGCGCTCAGACCG 540
Db 481 GCAGCTCATCGCTACAGCCCAACCTGGGCAAGTCTCATCTGCGAGCGCTCAGACCG 540
QY 541 CAGTCTACAGGCTTCCCGAGCGCCAGCACTGTGTGCAAGCAGGCTTTATGACTCGG 600
Db 541 CAGTCTACAGGCTTCCCGAGCGCCAGCACTGTGTGCAAGCAGGCTTTATGACTCGG 600
QY 601 CAGAACACAAAGCTTGGAGAACTGAGCCCTACACACAAATGTTCTGGGTCGCTGGG 660
Db 601 CAGAACACAAAGCTTGGAGAACTGAGCCCTACACACAAATGTTCTGGGTCGCTGGG 660
QY 661 TGTGGTTTCCCACTGTCAATGAAGGCTGTGGAGGTGCAATCCGGGACCTTATCC 720
Db 661 TGTGGTTTCCCACTGTCAATGAAGGCTGTGGAGGTGCAATCCGGGACCTTATCC 720
QY 721 TGTCTCAGAGCTGCTGTAAGAGATGAACACTTTCGTACTCACTGTGAGCACTGTATG 780
Db 721 TGTCTCAGAGCTGCTGTAAGAGATGAACACTTTCGTACTCACTGTGAGCACTGTATG 780
QY 781 CCTACGACTGATAGTATCCCACTGGTATTAACAGGTGTGACCTGTGGCGGTGAC 840
Db 781 CCTACGACTGATAGTATCCCACTGGTATTAACAGGTGTGACCTGTGGCGGTGAC 840
QY 841 GCTTCTTCTGACTTGTCTAGTGTGGGCGGAGTTTCTGAACCGAGCAAGGCTTACCTG 900
Db 841 GCTTCTTCTGACTTGTCTAGTGTGGGCGGAGTTTCTGAACCGAGCAAGGCTTACCTG 900
```

```
Db 841 GCTTCTTCTGACTTGTCTAGTGTGGGCGGAGTTTCTGAACCGAGCAAGGCTTACCTG 900
QY 901 GCCATGAGCTGAGCTCGTGTGCGCCGCTTACAGTCTCTGAGTCTTCTTATGTTG 960
Db 901 GCCATGAGCTGAGCTCGTGTGCGCCGCTTACAGTCTCTGAGTCTTCTTATGTTG 960
QY 961 GCTGCTTAAGTGTGACAGACGCTCAACCCCTTTGGAGAGATGATGATATTTTG 1020
Db 961 GCTGCTTAAGTGTGACAGACGCTCAACCCCTTTGGAGAGATGATGATATTTTG 1020
QY 1021 AGACCACTGATGATGTCAGACAGATTTGACAGTGTCCGTTGGCTGGATGATGATG 1080
Db 1021 AGACCACTGATGATGTCAGACAGATTTGACAGTGTCCGTTGGCTGGATGATGATG 1080
QY 1081 ACCAGACCTGCTCGATGAGGCGGACATGATCTGAAATTAAGCCGAGCAGAGCC 1140
Db 1081 ACCAGACCTGCTCGATGAGGCGGACATGATCTGAAATTAAGCCGAGCAGAGCC 1140
QY 1141 CCTACAGAGCTTCCGCGCCAGTTCGATGAGCTCTTTATGGCTCCACCTTCAACA 1200
Db 1141 CCTACAGAGCTTCCGCGCCAGTTCGATGAGCTCTTTATGGCTCCACCTTCAACA 1200
QY 1201 TCAGCTGAAACAAAGAGATGAGTTCAGCCCAATCAGAGGACAGAGAGATGCTC 1260
Db 1201 TCAGCTGAAACAAAGAGATGAGTTCAGCCCAATCAGAGGACAGAGAGATGCTC 1260
QY 1261 ACGTGGCATCATTTGGCGCTTCTTGAAGCTGCAATGCCATGATCAATCTCCAGAG 1320
Db 1261 ACGTGGCATCATTTGGCGCTTCTTGAAGCTGCAATGCCATGATCAATCTCCAGAG 1320
QY 1321 CAAACTCAAGACCAAACTACTGAGGCCCAAGAGGAAATCCCTTCCACAGAGGCTGC 1380
Db 1321 CAAACTCAAGACCAAACTACTGAGGCCCAAGAGGAAATCCCTTCCACAGAGGCTGC 1380
QY 1381 CCAAAAACCAAGGACGCCAAACAGAACGTTAGGGGCGAGAGAGCAACAGGCTTGA 1440
Db 1381 CCAAAAACCAAGGACGCCAAACAGAACGTTAGGGGCGAGAGAGCAACAGGCTTGA 1440
QY 1441 AGCTTAAAGCTGTGAGCGCTTCAAGTCTGAGCCCATCTGATCAGAGGCCAGCTTACA 1500
Db 1441 AGCTTAAAGCTGTGAGCGCTTCAAGTCTGAGCCCATCTGATCAGAGGCCAGCTTACA 1500
QY 1501 GTGCCCCACAGAGCGCCCTCAGGCCCATCCCATGTTCTTCCCTGAGAACCATGACGC 1560
Db 1501 GTGCCCCACAGAGCGCCCTCAGGCCCATCCCATGTTCTTCCCTGAGAACCATGACGC 1560
QY 1561 CTTCAAAAGCTTCAAGTGTCAAGGCAATGACACCAAGACAAAGCTTAAAGCTGTGA 1620
Db 1561 CTTCAAAAGCTTCAAGTGTCAAGGCAATGACACCAAGACAAAGCTTAAAGCTGTGA 1620
QY 1621 GTTCTGGGGCCAAAGAAATGTTTGAATGCTCTCAGAGAGGAGATGGGCTTGAATGAG 1680
Db 1621 GTTCTGGGGCCAAAGAAATGTTTGAATGCTCTCAGAGAGGAGATGGGCTTGAATGAG 1680
QY 1681 ACCCAGAAATATCTCAAGTGAAGAGAAATGTGGAATTTAACTGACGATATGCGAG 1740
Db 1681 ACCCAGAAATATCTCAAGTGAAGAGAAATGTGGAATTTAACTGACGATATGCGAG 1740
QY 1741 AGATCCCGGAAATATCACTCAAAAGAACTTTTGAATTCACCAACCAATACACTA 1800
Db 1741 AGATCCCGGAAATATCACTCAAAAGAACTTTTGAATTCACCAACCAATACACTA 1800
QY 1801 CACTCAAGATCATATGATCTTATTTGGGCTTGTGAAACAGGATTAAGCAATTTCT 1860
Db 1801 CACTCAAGATCATATGATCTTATTTGGGCTTGTGAAACAGGATTAAGCAATTTCT 1860
QY 1861 AACCTGCTTCTTAATGGGATGCTTTCGACAGGCTCTCACTGTGTGATACACAGCA 1920
Db 1861 AACCTGCTTCTTAATGGGATGCTTTCGACAGGCTCTCACTGTGTGATACACAGCA 1920
QY 1921 GGACACTGATCCAGTCAAGGCAATCAGCTGTCAACTGAAGAACTGTCTTCAACAG 1980
Db 1921 GGACACTGATCCAGTCAAGGCAATCAGCTGTCAACTGAAGAACTGTCTTCAACAG 1980
```

Qy 1981 CCTGAATCAATGTTAGCTTAATAGATTAATAATCCAGACTACTGAGCTTTAATGCC 2040
Db 1981 CCGAATCAAAATGTTAGCTTAATAGATTAATAATCCAGACTACTGAGCTTTAATGCC 2040
Qy 2041 TTTTATTCATTAATAAAGCTGTAAGCTAGACTGAACCATTTGAAAATTAACTCAGACTC 2100
Db 2041 TTTTATTCATTAATAAAGCTGTAAGCTAGACTGAACCATTTGAAAATTAACTCAGACTC 2100
Qy 2101 TGAATTCAGAGTGGGAAACCTTGTCTATCTGAATCCAGACGACACCTTAGTAT 2160
Db 2101 TGAATTCAGAGTGGGAAACCTTGTCTATCTGAATCCAGACGACACCTTAGTAT 2160
Qy 2161 ACTGCCCAACTAATGATTAAATAATACAAATATCTGTTAAAAAATTTTTAAAAA 2220
Db 2161 ACTGCCCAACTAATGATTAAATAATACAAATATCTGTTAAAAAATTTTTAAAAA 2220
Qy 2221 AAAAAAAAA 2229
Db 2221 AAAAAAAAA 2229

RESULT 2
AA221228
ID AA221228 standard; cDNA; 2429 BP.

AA221228;

22-NOV-1999 (first entry)

Human CG1CE long form cDNA sequence.

CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
age-related macular dystrophy; ss.

Homo sapiens.

Location/Qualifiers

Key CD5 /tag= a /product= "CG1CE short form protein"

MO9943695-A1.

02-SEP-1999.

22-FEB-1999; 99WO-US003790.

25-FEB-1998; 98US-0075941P.

18-DEC-1998; 98US-0112926P.

(MERI) MERCK & CO INC.

(UYUP-) UNIV UPPSALA.

Petrushin K, Caskey CT, Metzker M, Wadelius C;

WPI; 1999-540560/45.

P-PSDB; AAT2954.

Human and mouse polynucleotides encoding CG1CE polypeptides.

Claim 2; Fig 4; 67pp; English.

The present sequence represents the human CG1CE cDNA sequence, which when mutated is responsible for Best's macular dystrophy (BMD). Polynucleotides encoding CG1CE are useful for diagnosing whether a patient carries a mutation in the CG1CE gene. Normal and mutated CG1CE proteins are useful for identifying activators and/or inhibitors of these proteins, in order to treat BMD. The CG1CE gene offers a simpler and cheaper method of diagnosing BMD without the need for the presence of the patient. The gene may also be useful to discovering the genetic cause of age-related macular dystrophy

Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 U; 0 Other;

Query Match 90.3%; Score 2013; DB 2; Length 2429;

Best Local Similarity 91.6%; Pred. No. 0; Mismatches 0; Indels 203; Gaps 1;

Matches 2226; Conservative 0; Mismatches 0; Indels 203; Gaps 1;

Qy 1 CAGGAGTCCCAACAGCTTATGCGCAAGACCTTGTGGATCATCGGACCCACCTGGAA 60
Db 1 CAGGAGTCCCAACAGCTTATGCGCAAGACCTTGTGGATCATCGGACCCACCTGGAA 60
Qy 61 CCCCACTGACCCCAAGCCCACTGCTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTG 120
Db 61 CCCCACTGACCCCAAGCCCACTGCTGAGCCCACTGCTGCTGCTGCTGCTGCTGCTG 120
Qy 121 CAAGCCAAAGTGGCTTAATGCGGCTTAAAGCTCTCTCCGCGCTGCTGCTGCTGCTG 180
Db 121 CAAGCCAAAGTGGCTTAATGCGGCTTAAAGCTCTCTCCGCGCTGCTGCTGCTGCTG 180
Qy 181 GCAGCATCTCAAGCTGCTATATGCGAGTTCCTTAATCTTCTGCTGCTGCTGCTGCTA 240
Db 181 GCAGCATCTCAAGCTGCTATATGCGAGTTCCTTAATCTTCTGCTGCTGCTGCTGCTA 240
Qy 241 TCCGCTTTATTTATAGCTGCGCTTCAAGGAAACAACAGCTATGTTGAGAAACTGA 300
Db 241 TCCGCTTTATTTATAGCTGCGCTTCAAGGAAACAACAGCTATGTTGAGAAACTGA 300
Qy 301 CTCTGTATGGAAGCTATCAACGCTCATCCCATTTCTTGTGTGGGCTTCTACG 360
Db 301 CTCTGTATGGAAGCTATCAACGCTCATCCCATTTCTTGTGTGGGCTTCTACG 360
Qy 361 TGAAGCTGTCGTGACCCGCTGCTGGAACAAGTACGAGAACTGCGTGGCCGACCGCC 420
Db 361 TGAAGCTGTCGTGACCCGCTGCTGGAACAAGTACGAGAACTGCGTGGCCGACCGCC 420
Qy 421 TCATGAGCTGCTGCTGCGGCTTGTGGAAGCAAGACGAGCAAGCGGCTGCTGCGGC 480
Db 421 TCATGAGCTGCTGCTGCGGCTTGTGGAAGCAAGACGAGCAAGCGGCTGCTGCGGC 480
Qy 481 GCAGGCTCATCCGCTAGCGCAACCTGGCAACGCTCATCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GCAGGCTCATCCGCTAGCGCAACCTGGCAACGCTCATCTGCTGCTGCTGCTGCTGCTG 540
Qy 541 CAGTCTCAAGCGCTTCCCAAGCCGACGACCTGCTGCAAGCAAGCGCTTATGACTCGG 600
Db 541 CAGTCTCAAGCGCTTCCCAAGCCGACGACCTGCTGCAAGCAAGCGCTTATGACTCGG 600
Qy 601 CAGAACCAAGCAAGCTGGAAGAACTGACCTACCAACAATGTTCTGCGTGCCTGGG 660
Db 601 CAGAACCAAGCAAGCTGGAAGAACTGACCTACCAACAATGTTCTGCGTGCCTGGG 660
Qy 661 TGTGCTTTGCAACCTGCTCAATGAAAGGCGTGTGGAAGTGGCAATCCGGGACCTTATCC 720
Db 661 TGTGCTTTGCAACCTGCTCAATGAAAGGCGTGTGGAAGTGGCAATCCGGGACCTTATCC 720
Qy 721 TGTCCAGAGCTGCTGAAGAGATGAAACCTTGCCTACTCAGTGTGCAACCTGTATG 780
Db 721 TGTCCAGAGCTGCTGAAGAGATGAAACCTTGCCTACTCAGTGTGCAACCTGTATG 780
Qy 781 CTTAAGACTGATTAATATCCCACTGCTGTATACAGAGTGTGACTGTGCGGTGTACA 840
Db 781 CTTAAGACTGATTAATATCCCACTGCTGTATACAGAGTGTGACTGTGCGGTGTACA 840
Qy 841 GCTTCTTCCGACTTGTGCTTATGTTGGGGGAGTTCTGAAACCCGCAAGGCTTACCTG 900
Db 841 GCTTCTTCCGACTTGTGCTTATGTTGGGGGAGTTCTGAAACCCGCAAGGCTTACCTG 900
Qy 901 GCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 GCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Qy 961 GCTGCTGAA----- 970
Db 961 GCTGCTGAAAGTGGGCTCTCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

QY 971 ----- 970
Db 1021 ATGGCCAGCAGCTGTTGAGAGAGATGCAAGTTCAGAGAAAGAGTCTCACGGGTAG 1080
QY 971 ----- 970
Db 1081 AAGAGCAGCAGCGGTGTGGCGCACCTGTATCCAGCTACCTGGAGGCTGAGGCGAG 1140
QY 971 ----- GGTGGCAGAGCGAGCTCATACCCCTT 997
Db 1141 GAGAAATCGTTGAAACCCGGAGGCGGAGGTTGTGTGGCAGAGAGCTCATACCCCTT 1200
QY 998 TGGAGAGATGATGATGATTTTGAACAACCTGATTTGTGACAGGAAATTTGACAGTGTG 1057
Db 1201 TGGAGAGATGATGATGATTTTGAACAACCTGATTTGTGACAGGAAATTTGACAGTGTG 1260
QY 1058 CCGTGTGGCTGTGATGATGATGACACAGACCTGCTGTGGATGAGCGGACATGTACTG 1117
Db 1261 CCGTGTGGCTGTGATGATGATGACACAGACCTGCTGTGGATGAGCGGACATGTACTG 1320
QY 1118 GAATAGCCCGACAGCCAGCCCTTACACAGCTGTCCGCGCAGTTCCGTGAGCGCTC 1177
Db 1321 GAATAGCCCGACAGCCAGCCCTTACACAGCTGTCCGCGCAGTTCCGTGAGCGCTC 1380
QY 1178 CTTTATGGGCTTCACCTTCAACATCAGCTGAAACAAAGAGATGAGTTCCAGCCCA 1237
Db 1381 CTTTATGGGCTTCACCTTCAACATCAGCTGAAACAAAGAGATGAGTTCCAGCCCA 1440
QY 1238 TCAGAGAGCAGAGAGATGCTCAGGCTGGATCTATTGGCGCTTCCTAAGCTGCGAGTC 1297
Db 1441 TCAGAGAGCAGAGAGATGCTCAGGCTGGATCTATTGGCGCTTCCTAAGCTGCGAGTC 1500
QY 1298 CCATGATCACCCTCTCCAGAGGCAACTCAAGAGCAAACTAGTGTGGCCCAAGAGGA 1357
Db 1501 CCATGATCACCCTCTCCAGAGGCAAACTCAAGAGCAAACTAGTGTGGCCCAAGAGGA 1560
QY 1358 ATCCCTTTCACAGAGGCTGCTCCCAAAACCAAGGAGCCCAACAGAACTTTAGGGG 1417
Db 1561 ATCCCTTTCACAGAGGCTGCTCCCAAAACCAAGGAGCCCAACAGAACTTTAGGGG 1620
QY 1418 CCAGAGAGCAACAAGGCTGGAAGCTTAAAGCTGTGGAGGCTTCAAGTCTGGCCCACT 1477
Db 1621 CCAGAGAGCAACAAGGCTGGAAGCTTAAAGCTGTGGAGGCTTCAAGTCTGGCCCACT 1680
QY 1478 GTATCAGAGGCTCAGGCTTCAAGTGTCCCAACAGAGCCCTCAGCCCACTCCCATGTT 1537
Db 1681 GTATCAGAGGCTCAGGCTTCAAGTGTCCCAACAGAGCCCTCAGCCCACTCCCATGTT 1740
QY 1538 CTTCCCTTGAACCATCAGGCGCTCAAGCTTCAAGTGTCAAGGCTTCAAGGCTTCAAG 1597
Db 1741 CTTCCCTTGAACCATCAGGCGCTCAAGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 1800
QY 1598 AGACAAAGCTTAAAGCTGTGAGTCTGGGCGCAAGAAAGTTTGAATTTGCTCAGA 1657
Db 1801 AGACAAAGCTTAAAGCTGTGAGTCTGGGCGCAAGAAAGTTTGAATTTGCTCAGA 1860
QY 1658 GAGGAGTGGGCTTGAATGAGACCCCAAGATTTCTCAAGTGAAGAGAGAACTGTGA 1717
Db 1861 GAGGAGTGGGCTTGAATGAGACCCCAAGATTTCTCAAGTGAAGAGAGAACTGTGA 1920
QY 1718 GTTTAACTGACGGAATGTGCAGAGATCCCGAAATCACTCAAGAACTTTGGAACA 1777
Db 1921 GTTTAACTGACGGAATGTGCAGAGATCCCGAAATCACTCAAGAACTTTGGAACA 1980
QY 1778 ATCAACCAACAATACACTCACTCAAGATTCATGATCTTATTTGGGCTTTGA 1837
Db 1981 ATCAACCAACAATACACTCACTCAAGATTCATGATCTTATTTGGGCTTTGA 2040
QY 1838 AAACAGGAGTGAAGCATTTCTTAACCTGTTCTTATGGGAGTGTGGCCAGCAGGT 1897
Db 2041 AAACAGGAGTGAAGCATTTCTTAACCTGTTCTTATGGGAGTGTGGCCAGCAGGT 2100

QY 1898 CCTCACTGTGTGTACACACAGACAGACATGATCCAGTCAAGCCATACAGCTGTCCACA 1957
Db 2101 CCTCACTGTGTGTACACACAGACAGACATGATCCAGTCAAGCCATACAGCTGTCCACA 2160
QY 1958 CTGAAGAACGTGTCTCTCAACAGGCTGATCAAAATGTGTGCTTAAATGATTAATAATCCC 2017
Db 2161 CTGAAGAACGTGTCTCTCAACAGGCTGATCAAAATGTGTGCTTAAATGATTAATAATCCC 2220
QY 2018 AGACTACTTCAAGCTTTAATGCTTTTATTCATTAATACTGTGAAGCTAGTGAACCA 2077
Db 2221 AGACTACTTCAAGCTTTAATGCTTTTATTCATTAATACTGTGAAGCTAGTGAACCA 2280
QY 2078 TTGAAACATTTAATCTGAGCTGTGATTTAGAGTGGGAAACCTTAAGTTCTATGGAAT 2137
Db 2281 TTGAAACATTTAATCTGAGCTGTGATTTAGAGTGGGAAACCTTAAGTTCTATGGAAT 2340
QY 2138 CCAAGACGACACCTTATGATATGCTGCCCAATGATGATTTAAATCAAAATACT 2197
Db 2341 CCAAGACGACACCTTATGATATGCTGCCCAATGATGATTTAAATCAAAATACT 2400
QY 2198 CGTTAAAAA 2226
Db 2401 CGTTAAAAA 2429

RESULT 3
AD084435
ID AD084435 standard, cDNA, 2404 BP.
XX
AC AD084435;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1249.
XX
KW human; tumour-associated antigenic target; TAT; cyrostatic; gene therapy;
KM cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN NC02004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003W0-US029126.
PR 18-OCT-2002; 2002US-0418988P.
XX
PA (GENE) GENENTECH INC.
PA (WU/D/) WU T D.
PA (ZHOU/) ZHOU Y.
PI Wu TD, Zhou Y;
XX
DR WPI, 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
PS Claim 1; SEQ ID NO 1249; 5504bp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a)-(c); or (e) a sequence that hybridizes to (a)-
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide

Db	1681	GTATCAGAGGCCAGGTTACTTAAGTGGCCCCACAGAGCCCTTCAAGCCCATCTCCATGTT	1740
Qy	1538	CTTCCCCCTAGAAACCATCAGCGCCCGTCAAAAGCTTCAAGATGTCACAGGATAGACCA	1597
Db	1741	CTTCCCCCTAGAAACCATCAGCGCCCGTCAAAAGCTTCAAGATGTCACAGGATAGACCA	1800
Qy	1598	AGACAAAGCTTAAAGACTGTAGTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGA	1657
Db	1801	AGACAAAGCTTAAAGACTGTAGTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGA	1866
Qy	1658	GAGCATGGGAGCTTGGATGGAGCAACCAAGATCTCAGTGAAGAGAAAGCTGTGGA	1717
Db	1861	GAGCATGGGAGCTTGGATGGAGCAACCAAGATCTCAGTGAAGAGAAAGCTGTGGA	1920
Qy	1718	GTTTAACCTGACGGATATGCCAGATATCCCGAAATCACTCAAGAACCTTTGGAAACA	1777
Db	1921	GTTTAACCTGACGGATATGCCAGATATCCCGAAATCACTCAAGAACCTTTGGAAACA	1980
Qy	1778	ATCACCAACCAACATACACACTACACTCAAGATCAGATGGATCCTTATGGGCTTGGGA	1837
Db	1981	ATCACCAACCAACATACACACTACACTCAAGATCAGATGGATCCTTATGGGCTTGGGA	2040
Qy	1838	AAACGAGATGAAGCAATTCCTTAACCTGCTTCTTAATGGGAGATGCTTGGCAGCAGGT	1897
Db	2041	AAACGAGATGAAGCAATTCCTTAACCTGCTTCTTAATGGGAGATGCTTGGCAGCAGGT	2100
Qy	1898	CCTCACTGTGTGTACACAGAGAGCACTGATCCAGTCAAGGCATACAGCTGTCCACA	1957
Db	2101	CCTCACTGTGTGTACACAGAGAGCACTGATCCAGTCAAGGCATACAGCTGTCCACA	2160
Qy	1958	CTGAAGACGTGTCTCTACAAACAGCCTGATCAAAATGTTAGCTTAATAGATAAAGATCCC	2017
Db	2161	CTGAAGACGTGTCTCTACAAACAGCCTGATCAAAATGTTAGCTTAATAGATAAAGATCCC	2220
Qy	2018	AGACTACTTCAAGCCTTTAATGCTTTTATCATAAAGCTGTGAAGCTAGACTGAACCA	2077
Db	2221	AGACTACTTCAAGCCTTTAATGCTTTTATCATAAAGCTGTGAAGCTAGACTGAACCA	2280
Qy	2078	TTGGAAACATTTAACTCAGACTCTGAGATTGAGATCGGAAACCTTATGTTCTATCTGAAT	2137
Db	2281	TTGGAAACATTTAACTCAGACTCTGAGATTGAGATCGGAAACCTTATGTTCTATCTGAAT	2340
Qy	2138	CCAAAGACAGCCACACCTTAGTATCTGCCCCAAAGTAAATGAGTTTAAATACAAATACT	2197
Db	2341	CCAAAGACAGCCACACCTTAGTATCTGCCCCAAAGTAAATGAGTTTAAATACAAATACT	2400
Qy	2198	CGTT 2201	
Db	2401	CGTT 2404	
RESULT 4			
AD083262			
ID AD083262 standard; cDNA; 2404 BP.			
XX			
AC AD083262;			
XX			
DT 07-OCT-2004 (first entry)			
XX			
DE Human tumour-associated antigenic target (TAT) cDNA sequence #76.			
XX			
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;			
XX			
OS cancer; cell proliferative disorder; gene; ss.			
XX			
PN Homo sapiens.			
XX			
PD WO2004060270-A2.			
XX			
PF 22-JUL-2004.			
XX			
15-OCT-2003; 2003WO-US029126.			
XX			

18-OCT-2002; 2002US-0418988P.

(GETH) GENENTECH INC.
(WUTD/) WU T D.
(ZHOU/) ZHOU Y.

Wu TD, Zhou Y;
WPI, 2004-534300/51.

New nucleic acid molecule and encoded polypeptide, for diagnosing, preventing or treating cell proliferative disorders such as cancer.

Claim 1; SEQ ID NO 76; 5504bp; English.

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentially effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumour comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing a medicament for the therapeutic treatment or diagnostic detection of a cell proliferative disorder or cancer. The present sequence represents a human TAT cDNA sequence from the present invention.

Sequence 2404 BP; 590 A; 696 C; 611 G; 507 T; 0 U; 0 Other;

Query Match 89.0%; Score 1984.8; DB 13; Length 2404;
Best Local Similarity 91.5%; Prid. No. 0;
Matches 2199; Conservative 0; Mismatches 2; Indels 203; Gaps 14

1 CAGGAGATGCCACAGCCTAGTGGCCAGACCTTCGTGGGATCATCGGACCACTGGAA 60
1 CAGGAGATGCCACAGCCTAGTGGCCAGACCTTCGTGGGATCATCGGACCACTGGAA 60

61 CCCCACTGACCCAGGCCACTGCTCAGGCCACTGCGCCATGACATCACTTACA 120
61 CCCCACTGACCCAGGCCACTGCTCAGGCCACTGCGCCATGACATCACTTACA 120

121 CAAGCCAAAGTGGCTAATGGCCCGCTTAGGCTCTTCGCCGCTGCTGTCGCGCGG 180
121 CAAGCCAAAGTGGCTAATGGCCCGCTTAGGCTCTTCGCCGCTGCTGTCGCGCGG 180

181 GAGCAGCTTACAGGCTGCTATATGGCAGATGTTTAACTCTTGGCTCTGCTACTACATCA 240

```
Db 181 |GCACATCTACAACTGCTATATGCGAGTCTTAATCTTCTGCTGCTACTACATCA 240
Qy 241 |TCCGCTTATTTATATAGCTGGCCCTCAGAGAAACAACAGCTATGTTTGAAGAACTGA 300
Db 241 |TCCCTTTATTTATATAGCTGGCCCTCAGAGAAACAACAGCTATGTTTGAAGAACTGA 300
Qy 301 |CTCTGTATTTGGACAGCTACATCCAGCTCATCCCATTTCTTGCTGGGCTTTCTACG 360
Db 301 |CTCTGTATTTGGACAGCTACATCCAGCTCATCCCATTTCTTGCTGGGCTTTCTACG 360
Qy 361 |TGACGCTGTGTGTGACCCGCTGTGTGAAACAAGTACAGAACCTGCGTGCACGCGC 420
Db 361 |TGACGCTGTGTGTGACCCGCTGTGTGAAACAAGTACAGAACCTGCGTGCACGCGC 420
Qy 421 |TCATGACCTGTGTGTGCTGGGCTTTGTCGAAGGCAAGCAAGGCGCGCTGTGCTGGC 480
Db 421 |TCATGACCTGTGTGTGCTGGGCTTTGTCGAAGGCAAGCAAGGCGCGCTGTGCTGGC 480
Qy 481 |GCACGCTCATCCGCTACGCCAACCTGGGCAACGTGCTATCTGTGCGAGCGTCAAGACCG 540
Db 481 |GCACGCTCATCCGCTACGCCAACCTGGGCAACGTGCTATCTGTGCGAGCGTCAAGACCG 540
Qy 541 |CAGTCTACAAAGCGCTTCCCAAGCGCCCAAGCCTGTGTGCAAGAGGCTTTATGACTCCGG 600
Db 541 |CAGTCTACAAAGCGCTTCCCAAGCGCCCAAGCCTGTGTGCAAGAGGCTTTATGACTCCGG 600
Qy 601 |CAGAAACACAAAGCAGTTGGAGAAATGAGCTTACCAACATGTTCTGGGTGCTGGG 660
Db 601 |CAGAAACACAAAGCAGTTGGAGAAATGAGCTTACCAACATGTTCTGGGTGCTGGG 660
Qy 661 |TGTGTTTTGGCAACCTGTCAATGAAAGGCTGTGTGAGGTGGAATCCGGGACCTTATCC 720
Db 661 |TGTGTTTTGGCAACCTGTCAATGAAAGGCTGTGTGAGGTGGAATCCGGGACCTTATCC 720
Qy 721 |TGCTCCAGAGCCTGCTGAACGAGATGAAACACCTTGCGTACTCAGTGTGACACCTGTATG 780
Db 721 |TGCTCCAGAGCCTGCTGAACGAGATGAAACACCTTGCGTACTCAGTGTGACACCTGTATG 780
Qy 781 |CCTACGACTGGATTAATATCCCACTGTGTATTAACAGAGTGTGACTGTGGCGGTATCA 840
Db 781 |CCTACGACTGGATTAATATCCCACTGTGTATTAACAGAGTGTGACTGTGGCGGTATCA 840
Qy 841 |GCTTCTTCTGACTGTGTAGTTGGGGGGGAGTTTCTGAACCCAGCCAGGCTTATACCTG 900
Db 841 |GCTTCTTCTGACTGTGTAGTTGGGGGGGAGTTTCTGAACCCAGCCAGGCTTATACCTG 900
Qy 901 |GCCATGAGCTGGACCTGCTGTGTGCGGCTTCAAGTTGCTGTGCAAGTTCTTCTATGTTG 960
Db 901 |GCCATGAGCTGGACCTGCTGTGTGCGGCTTCAAGTTGCTGTGCAAGTTCTTCTATGTTG 960
Qy 961 |GCTGGCTGGA----- 970
Db 961 |GCTGGCTGGAAGTGGGCTCTTCCAGGGCCCTGTGGGCTGAGGCAATG3CCAAGAGGGTTC 1020
Qy 971 |----- 970
Db 971 |ATGGCCACAGCTGCTGAGACAGAGATGCAAGTCAAGAAAAGAAAGTCTTCAAGGGTAG 1080
Qy 971 |----- 970
Db 971 |----- 970
Qy 971 |-----GTTGGCAGAGCAGCTCATCAACCCCTT 997
Db 971 |-----GTTGGCAGAGCAGCTCATCAACCCCTT 997
Qy 998 |TGGAGAGGATGATGATTTTGAAGCAACCTGATTTGAGAGAGAAATTTGAGAGGTTC 1057
Db 998 |TGGAGAGGATGATGATTTTGAAGCAACCTGATTTGAGAGAGAAATTTGAGAGGTTC 1057
Qy 1201 |TGGAGAGGATGATGATTTTGAAGCAACCTGATTTGAGAGAGAAATTTGAGAGGTTC 1260
Db 1201 |TGGAGAGGATGATGATTTTGAAGCAACCTGATTTGAGAGAGAAATTTGAGAGGTTC 1260
Qy 1058 |CCTGTGGCTGTGATGATGACACAGACCTGCTCGATGAGCGCGGACATGTACTG 1117
Db 1058 |CCTGTGGCTGTGATGATGACACAGACCTGCTCGATGAGCGCGGACATGTACTG 1117
Qy 1261 |CCTGTGGCTGTGATGATGACACAGACCTGCTCGATGAGCGCGGACATGTACTG 1320
Qy 1118 |GAATTAAGCCGAGGACACAGGCCCTTCAACAGCTGCTTCCGCGGAGTTCCGTGAGCTTC 1177
Db 1321 |GAATTAAGCCGAGGACACAGGCCCTTCAACAGCTGCTTCCGCGGAGTTCCGTGAGCTTC 1380
Qy 1178 |CTTTATGGGCTTCACTTCAACATCAAGCTTGAACAAAGAGAGATGAGATTCCAGCCCA 1237
Db 1381 |CTTTATGGGCTTCACTTCAACATCAAGCTTGAACAAAGAGAGATGAGATTCCAGCCCA 1440
Qy 1238 |TCAGAGAGACGAGAGAGATGCTCAAGCTGTGATCATTTGGCGCTTCTTAGGCTGCACTC 1297
Db 1441 |TCAGAGAGACGAGAGAGATGCTCAAGCTGTGATCATTTGGCGCTTCTTAGGCTGCACTC 1500
Qy 1298 |CCATGATCAACCATCTCCAGGGGCAACTCAAGAGCCAAATCACTGTGCGCCCAAGAGGA 1357
Db 1501 |CCATGATCAACCATCTCCAGGGGCAACTCAAGAGCCAAATCACTGTGCGCCCAAGAGGA 1560
Qy 1358 |ATCCCTTCTCCACGAGGGCCCTGCCAAAACCAAGAGCAGCCAAACGTTAGGGG 1417
Db 1561 |ATCCCTTCTCCACGAGGGCCCTGCCAAAACCAAGAGCAGCCAAACGTTAGGGG 1620
Qy 1418 |CCAGAAAGACAAAGGCTGTGAAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGGCCCACT 1477
Db 1621 |CCAGAAAGACAAAGGCTGTGAAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGGCCCACT 1680
Qy 1478 |GTATCAGAGGCAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTT 1537
Db 1681 |GTATCAGAGGCAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTT 1740
Qy 1538 |CTTCCCTTGAAGACCATCAAGGCGCTGAAAGCTTCAAGTGTCAAGGATGACACCA 1597
Db 1741 |CTTCCCTTGAAGACCATCAAGGCGCTGAAAGCTTCAAGTGTCAAGGATGACACCA 1800
Qy 1598 |AGACAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTT 1657
Db 1801 |AGACAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTT 1860
Qy 1658 |GAGGATGAGGGCTTGTGATGAGGACCCAGAAAGTATCTCAAGTGAAGGAGAAACGTGGA 1717
Db 1861 |GAGGATGAGGGCTTGTGATGAGGACCCAGAAAGTATCTCAAGTGAAGGAGAAACGTGGA 1920
Qy 1718 |GTTTAACCTGACGAGATATGACAGAGATCCCGGAAATCACTCAAGAACTTTGAAACA 1777
Db 1921 |GTTTAACCTGACGAGATATGACAGAGATCCCGGAAATCACTCAAGAACTTTGAAACA 1980
Qy 1778 |ATCACCAACCAACATACACATCACTCAAGATCAATGATCTTATTTGGGCTTGGGA 1837
Db 1981 |ATCACCAACCAACATACACATCACTCAAGATCAATGATCTTATTTGGGCTTGGGA 2040
Qy 1838 |AAAAGAGGATGAGACACATTTCTTAACTGTCTTCTTAAATGGGAGATGCTTGGCAGCGAGT 1897
Db 2041 |AAAAGAGGATGAGACACATTTCTTAACTGTCTTCTTAAATGGGAGATGCTTGGCAGCGAGT 2100
Qy 1898 |CCTCACTGTGTGACACACAGAGAGACATGATCAAGTCAAGGACATACAGCTGTCCACA 1957
Db 2101 |CCTCACTGTGTGACACACAGAGAGACATGATCAAGTCAAGGACATACAGCTGTCCACA 2160
Qy 1958 |CTGAAGAACGTTGCTTAAACAGGCTGAAATCAAAATGTTAGCTTAAATGATTAATAAATCC 2017
Db 2161 |CTGAAGAACATGTTCTTAAACAGGCTGAAATCAAAATGTTAGCTTAAATGATTAATAAATCC 2220
Qy 2018 |AGACTACTTCAAGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATG 2077
Db 2221 |AGACTACTTCAAGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATG 2280
Qy 2078 |TTGGAACATTTAACTCAGACTGTGATTCAGATTCGGGAAACCTTATGTTCTATCTGAT 2137
Db 2281 |TTGGAACATTTAACTCAGACTGTGATTCAGATTCGGGAAACCTTATGTTCTATCTGAT 2340
Qy 2138 |CAAAGACGACACCTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2197
Db 2341 |CAAAGACGACACCTTAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
```

D	b	1	ATGACATCACTTACACAAAGCAAGTGGCTTAATGCCGCTTAAGCTCCTCTTCCGCTTG	60
Q	y	165	CTGCTGTGCTGCGGGGAGCATCTACAGCTGCTTAATGGCGAGTTCTTAATCTTCCTG	22
D	b	61	CTGCTGTGCTGCGGGGAGCATCTACAGCTGCTTAATGGCGAGTTCTTAATCTTCCTG	122
Q	y	225	CTGCTGCTACATCAATCCGCTTAATTATATAGCTGGCCCTCAACGAAACAACAGCTG	284
D	b	121	CTTGCTACTACATCAATCCGCTTAATTATATAGCTGGCCCTCAACGAAACAACAGCTG	188
Q	y	285	ATGTTTGAAGAACTGACTCTGTATTGGACAGACTACATCCAGCTCAATCCCATTTCCCTC	344
D	b	181	ATGTTTGAAGAACTGACTCTGTATTGGACAGACTACATCCAGCTCAATCCCATTTCCCTC	240
Q	y	345	GTGTGTGGGCTTTTACATGACAGCTGTGTGTGACCCGCTGTGGAACAGTACGAAGACTTG	404
D	b	241	GTGTGTGGGCTTTTACATGACAGCTGTGTGTGACCCGCTGTGGAACAGTACGAAGACTTG	300
Q	y	405	CCGTGAGCCACACGCGCTCATGAGCTGTGTGTGGGCTTTGTGCAAGGCAAGACGAGCA	464
D	b	301	CCGTGAGCCACACGCGCTCATGAGCTGTGTGTGGGCTTTGTGCAAGGCAAGACGAGCA	360
Q	y	465	GCGCGGCTGTGCGCGCGACGCTCATCGCTACGCCAACCTGGGCGAACGCTCATCTTG	524
D	b	361	GCGCGGCTGTGCGCGCGACGCTCATCGCTACGCCAACCTGGGCGAACGCTCATCTTG	420
Q	y	525	CGAGAGGTGACACCGCAGCTTACAAAGGCGCTTCCACAGCGCCAGCATCTGGTGCAAGCA	584
D	b	421	CGAGAGGTGACACCGCAGCTTACAAAGGCGCTTCCACAGCGCCAGCATCTGGTGCAAGCA	480
Q	y	585	GCGCTTATATACCTCCGACAGAACACAGAGTTGGAGAACTGAGCTTACACACAAACATG	644
D	b	481	GCGCTTATATACCTCCGACAGAACACAGAGTTGGAGAACTGAGCTTACACACAAACATG	540
Q	y	645	TTCTGGGTGCGCTGGGTGTGGTTTGGCAACTGTATATGAAGGCGTGGCTTGGAGTGTGA	704
D	b	541	TTCTGGGTGCGCTGGGTGTGGTTTGGCAACTGTATATGAAGGCGTGGCTTGGAGTGTGA	600
Q	y	705	ATCCGAGAACCTTATCTGCTCCAGAGCTGCTGAAACGATGAAGAACACTTGTGCTATAG	764
D	b	601	ATCCGAGAACCTTATCTGCTCCAGAGCTGCTGAAACGATGAAGAACACTTGTGCTATAG	660
Q	y	765	TGTGACACCTGTATGCTTACGACTGTGATTAATCCACTGTGTATATACACAGTGTGTG	824
D	b	661	TGTGACACCTGTATGCTTACGACTGTGATTAATCCACTGTGTATATACACAGTGTGTG	720
Q	y	825	ACTGAGCGGAGTGTACAGCTTCTTCTCTACATTTGTATTTGGGCGGCACTTCTGAACCA	884
D	b	721	ACTGAGCGGAGTGTACAGCTTCTTCTCTACATTTGTATTTGGGCGGCACTTCTGAACCA	780
Q	y	885	GCCAAAGGCTTACCTTGCGCATGAGCTGAACTCTGTTGCGCGTCTCAAGTTCCTGACG	944
D	b	781	GCCAAAGGCTTACCTTGCGCATGAGCTGAACTCTGTTGCGCGTCTCAAGTTCCTGACG	840
Q	y	945	TTCTTCTTATATGTTGGCTGCTGAAAGTGTGGCAGAGCAAGCTCATCAACCCCTTTGAGAG	1004
D	b	841	TTCTTCTTATATGTTGGCTGCTGAAAGTGTGGCAGAGCAAGCTCATCAACCCCTTTGAGAG	900
Q	y	1005	GATGATGATGATTTTGGACCAACTGGAATTTGTGACAGGAATTGGACAGGTGCCCTGTG	1064
D	b	901	GATGATGATGATTTTGGACCAACTGGAATTTGTGACAGGAATTGGACAGGTGCCCTGTG	960
Q	y	1065	GCCTGTGATATGATGACCAAGGACTTGCTCGATGAGACCGGACATGTACTGGAATTAAG	1124
D	b	961	GCCTGTGATATGATGACCAAGGACTTGCTCGATGAGACCGGACATGTACTGGAATTAAG	1020
Q	y	1125	CCCGAGCACAAGCCCCCTTACACAGCTGTCCGCCCACTTCCGTGAGACCTCTTTAAG	1184
D	b	1021	CCCGAGCACAAGCCCCCTTACACAGCTGTCCGCCCACTTCCGTGAGACCTCTTTAAG	1080
Q	y	1185	GCGTCAACTTAAACATCAAGCTTGAACAAAGAGAGATGAATTCAGCGCCATACAGAG	1244
D	b	1081	GCGTCAACTTAAACATCAAGCTTGAACAAAGAGAGATGAATTCAGCGCCATACAGAG	1140

QY 1245 GACGAGAGATGCTCAGCTGCGATCATTTGGCCGCTTCTAGGCTTCAGTCCCATGAT 1304
DB 1141 GACGAGAGATGCTCAGCTGCGATCATTTGGCCGCTTCTAGGCTTCAGTCCCATGAT 1200
QY 1305 CACCATCTCTCCAGGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTT 1364
DB 1201 CACCATCTCTCCAGGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTT 1260
QY 1365 CTCACGAGGGGCTGGCCCAAAACCAAGAGGAGCAAAAGCTTAAAGGGGCAAGAA 1424
DB 1261 CTCACGAGGGGCTGGCCCAAAACCAAGAGGAGCAAAAGCTTAAAGGGGCAAGAA 1320
QY 1425 GACAAACAGGCTCTGAAAGCTTAAAGCTGTGAGACGCTTCAAGTCTGAGCCCACTGATCAG 1484
DB 1321 GACAAACAGGCTCTGAAAGCTTAAAGCTGTGAGACGCTTCAAGTCTGAGCCCACTGATCAG 1380
QY 1485 AGGCGAGGCTACTACAGTGGCCCAAGAGGCGCCCTCAGGCCCACTCCCATGTTCTCCCC 1544
DB 1381 AGGCGAGGCTACTACAGTGGCCCAAGAGGCGCCCTCAGGCCCACTCCCATGTTCTCCCC 1440
QY 1545 CTAAACCATCAGGCGCGCTCAAGCTTCAAGTGTGACAGGATAGACCAAAAGACAA 1604
DB 1441 CTAAACCATCAGGCGCGCTCAAGCTTCAAGTGTGACAGGATAGACCAAAAGACAA 1500
QY 1605 AGCTTAAAGACTGTGAGTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCAT 1664
DB 1501 AGCTTAAAGACTGTGAGTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCAT 1560
QY 1665 GGGGCTCTTGTGAGACCAAGCAAGATATCTCAAGTGAAGAGAAAGCTGTGAAGTTAAC 1724
DB 1561 GGGGCTCTTGTGAGACCAAGCAAGATATCTCAAGTGAAGAGAAAGCTGTGAAGTTAAC 1620
QY 1725 CTGACGATATGCGAGAGATCCCGGAAATACCTCAAGAAAGAACTTTGGAACATTCACA 1784
DB 1621 CTGACGATATGCGAGAGATCCCGGAAATACCTCAAGAAAGAACTTTGGAACATTCACA 1680
QY 1785 ACCAATATACACTACACTCAAGATCATGATCTTATTTGGGCTTTGAAAGACAG 1844
DB 1681 ACCAATATACACTACACTCAAGATCATGATCTTATTTGGGCTTTGAAAGACAG 1740
QY 1845 GATGAAGCAATTCCTTA 1862
DB 1741 GATGAAGCAATTCCTTA 1758

RESULT 6
AAV9722
ID AAV9722 standard; cDNA; 1263 BP.
XX
AC AAV9722;
XX
DT 26-APR-1999 (first entry)
XX
DE Human adult retina secreted protein bk112_15 cDNA.
XX
KW Secreted protein; human; retina; bk112_15; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 100..885
FT /**tag= a
XX
FN MO9856909-A2.
XX
PD 17-DEC-1998.
XX
PE 08-JUN-1998; 98MO-US011822.
XX
PR 11-JUN-1997; 97US-00873218.
PR 05-JUN-1998; 98US-00092722.
XX

PA (GENY) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallee ER, Racine LA, Treacy M, Spaulding V;
PI Agostino MJ, Howes SH, Feichel K;
XX
DR WPI; 1999-080899/07.
XX
XX P-PSDB; AAW95345.
PT New polynucleotides encoding secreted human proteins - derived from human
PT foetal brain, adult testes, foetal kidney, adult thyroid or adult retina
PT cDNA libraries.
XX
XX
XX Claim 12; Page 71-72; 113pp; English.
XX
XX This is the nucleotide sequence of cDNA clone bk112_15, which includes an
XX open reading frame for a 261-amino acid polypeptide (see AAW95345). The
XX clone was isolated from a human adult retina cDNA library using methods
XX which are selective for cDNAs encoding secreted proteins, or was
XX identified as encoding a secreted or transmembrane protein on the basis
XX of computer analysis of the amino acid sequence of the encoding protein.
XX Database searches indicate some sequence similarity to known sequences.
XX The invention provides cDNA clones (see AAV9721-33) from human adult
XX thyroid, adult retina, adult testis, foetal kidney and foetal brain that
XX encode novel secreted proteins (see AAW95344-53). Each clone is
XX individually available from deposit clone ATCC 98451 (see also AAV9734-
XX 43). The isolated polynucleotides (PNS) and proteins are predicted to
XX have activities which would make them suitable for treating, preventing
XX or ameliorating medical conditions in humans and animals, although no
XX supporting data is given. Suggested activities include nutritional,
XX cytokine, cell proliferation or differentiation, immune stimulating (e.g.
XX as vaccines) or immune suppressing, haematopoietic regulating, tissue
XX growth, activin/inhibin, chemotactic/chemokinetic, haemostatic,
XX thrombolytic, receptor/ligand, anti-inflammatory, cadherin/tumour invasion
XX suppressor, and tumour inhibition activities. The PNS are also stated to
XX be useful for gene therapy
XX
XX
XX Sequence 1263 BP; 397 A; 342 C; 272 G; 252 T; 0 U; 0 Other;
XX
XX
XX Query Match 56.0%; Score 1247.2; DB 2; Length 1263;
XX Best Local Similarity 99.8%; Pred. No. 1e-280;
XX Matches 1249; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 978 GAGCAGCTCATCAACCCCTTTGAGAGATGATGATTTTGAAGCACTGATTTGTC 1037
DB 1 GAGCAGCTCATCAACCCCTTTGAGAGATGATGATTTTGAAGCACTGATTTGTC 60
QY 1038 GACAGAAATTTGACAGTGTCTCTGTTGCTGTGTGATGATGACCAAGAGCTGCTCGG 1097
DB 61 GACAGAAATTTGACAGTGTCTCTGTTGCTGTGTGATGATGACCAAGAGCTGCTCGG 120
QY 1098 ATGAGCCGGAACATGTAATGTAATTAAGCCGAGCAAGCCCTTACAGCTGCTTCC 1157
DB 121 ATGAGCCGGAACATGTAATGTAATTAAGCCGAGCAAGCCCTTACAGCTGCTTCC 180
QY 1158 GCCAGTTCCGTGAGGCTCTCTTATGAGGCTCCACTTCAACATCAGCTGAAAGAG 1217
DB 181 GCCAGTTCCGTGAGGCTCTCTTATGAGGCTCCACTTCAACATCAGCTGAAAGAG 240
QY 1218 GAGATGAGTTCCAGGCCAATCAGAGAGCAGAGAGATGCTACGCTGATATTGGC 1277
DB 241 GAGATGAGTTCCAGGCCAATCAGAGAGCAGAGAGATGCTACGCTGATATTGGC 300
QY 1278 CGCTTCTAAGGCTCGAGTGTCCATGATCAACATCTCCAGGGGAACTCAAGAGCA 1337
DB 301 CGCTTCTAAGGCTCGAGTGTCCATGATCAACATCTCCAGGGGAACTCAAGAGCA 360
QY 1338 CTACTGTGCCCCAAGAGGAATCCCTTCTCCACGAGGAGCTGCCCAAAACCAAGAGCA 1397
DB 361 CTACTGTGCCCCAAGAGGAATCCCTTCTCCACGAGGAGCTGCCCAAAACCAAGAGCA 420
QY 1398 GCCAAACAGAACTTTAGGGGCGAGGAAGACAAAGGCTTGAAGCTTAAAGCTGTGAGC 1457
DB 421 GCCAAACAGAACTTTAGGGGCGAGGAAGACAAAGGCTTGAAGCTTAAAGCTGTGAGC 480

QY 1458 GCGTTCAGTCTGCGCCACTGATCAGAGGCGGCTACTACAGTCCGCCACAGACGCC 1517
 DB 481 GCGTTCAGTCTGCGCCACTGATCAGAGGCGGCTACTACAGTCCGCCACAGACGCC 540
 QY 1518 CTCAGCCCCCACTCCCATGTTCTTCCCTAGAACCATCAGGCGCGTCAAGCTTCACAGT 1577
 DB 541 CTCAGCCCCCACTCCCATGTTCTTCCCTAGAACCATCAGGCGCGTCAAGCTTCACAGT 600
 QY 1578 GTCAAGGCAATAGACACCAAGAAAGAAAGCTTAAGACTGTGAGTTCTGGGCGCAAGAAA 1637
 DB 601 GTCAAGGCAATAGACACCAAGAAAGAAAGCTTAAGACTGTGAGTTCTGGGCGCAAGAAA 660
 QY 1638 AGTTTGAATTCCTCTCAGAGAGCGATGGGCGCTTGAAGAGACCCAGAAAGTATCTCAA 1697
 DB 661 AGTTTGAATTCCTCTCAGAGAGCGATGGGCGCTTGAAGAGACCCAGAAAGTATCTCAA 720
 QY 1698 GTGAGAGAGAAACCTGAGAGTTTAACCTGACGATATGCGAAGATCCCGAAATCAAC 1757
 DB 721 GTGAGAGAGAAACCTGAGAGTTTAACCTGACGATATGCGAAGATCCCGAAATCAAC 780
 QY 1758 CTCAAGAACCTTTGGAACATCAACCAACATACACACTACCTCAAAAGATCAATG 1817
 DB 781 CTCAAGAACCTTTGGAACATCAACCAACATACACACTACCTCAAAAGATCAATG 840
 QY 1818 GATCCTTAATGGGCGCTTGGAAACAGGAGTGAAGCACTTCTTAACCTGCTTCTAATG 1877
 DB 841 GATCCTTAATGGGCGCTTGGAAACAGGAGTGAAGCACTTCTTAACCTGCTTCTAATG 900
 QY 1878 GATGCTTGGCGAGCGAGTCTCACCTGTGTGTACACAGAGAGACACTGATCCAGTCA 1937
 DB 901 GATGCTTGGCGAGCGAGTCTCACCTGTGTGTACACAGAGAGACACTGATCCAGTCA 960
 QY 1938 CAGCCTAAGAGCTGTCCACACTGAGAAAGAGTGTCTCAACAGCCTGATCAATGTTA 1997
 DB 961 CAGCCTAAGAGCTGTCCACACTGAGAAAGAGTGTGTCTCAACAGCCTGATCAATGTTA 1020
 QY 1998 GCTTAATGATTAATAATCCAGACTACTCAGCCTTTATGCTTTATGTAATAAACT 2057
 DB 1021 GCTTAATGATTAATAATCCAGACTACTCAGCCTTTATGCTTTATGTAATAAACT 1080
 QY 2058 GTGAAGAGCTAGACTGAACATTTGGAAGCAATTTAATCAGACTGTGATTCGGGGA 2117
 DB 1081 GTGAAGAGCTAGACTGAACATTTGGAAGCAATTTAATCAGACTGTGATTCGGGGA 1140
 QY 2118 ACCCTTAATGCTATCTGAATCCAGAGAGCAACCTTAATGCTGCTCCAACTAATGA 2177
 DB 1141 ACCCTTAATGCTATCTGAATCCAGAGAGCAACCTTAATGCTGCTCCAACTAATGA 1200
 QY 2178 GTTAAATTAATTAATAATCTGTTAAAAAAGAAAAAAGAAAAA 2229
 DB 1201 GTTAAATTAATTAATAATCTGTTAAAAAAGAAAAAAGAAAAA 1252

RESULT 7
 ADA44960
 ID ADA44960 standard; cDNA; 1238 BP.
 AC ADA44960;
 DT 20-NOV-2003 (first entry)
 DE Human polynucleotide #2.
 XX Human; gene; ss; genetic disorder; genetic fingerprinting;
 KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
 KW insulin dependent diabetes mellitus; graft-versus-host disease; anaemia;
 KW periodontal disease; bone fracture; cartilage damage;
 KW central nervous system disorder; Alzheimer's disease;
 KW Parkinson's disease;
 KW carbohydrate source; cancer; nutrition; carbon source; nitrogen source;
 XX Homo sapiens.

XX US2003044935-A1.
 PN 06-MAR-2003.
 XX 21-DEC-2000; 2000US-00746783.
 XX 11-JUN-1997; 97US-0086234P.
 PR 12-JUN-1997; 97US-0086234P.
 PR 08-JUL-1997; 97US-0092115P.
 PR 08-SEP-1997; 97US-0093045P.
 PR 02-OCT-1997; 97US-0093045P.
 PR 27-OCT-1997; 97US-00958304.
 PR 07-NOV-1997; 97US-0090111P.
 PR 05-JUN-1998; 98US-00092722.
 PR 11-JUN-1998; 98US-00096287.
 PR 17-JUN-1998; 98US-00096287.
 PR 04-AUG-1998; 98US-00130189.
 PR 08-SEP-1998; 98US-00149633.
 PR 01-OCT-1998; 98US-00165960.
 PR 04-NOV-1998; 98US-00185936.
 XX (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J M.
 PA (LVAL/) LA VALLIE E R.
 PA (COLL/) COLLINS-RACIE L A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREA/) TREACY M.
 PA (SPAU/) SPAULDING V.
 PI Jacobs K, McCoy JM, La Vallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 PI WPI; 2003-521754/49.
 DR P-PSDB; ADA44961.
 XX New polypeptides and polynucleotides having biological activities, useful
 PT as nutritional sources or supplements, or for treating e.g. autoimmune
 PT diseases, cancers, bone fractures or damages, or central nervous system
 PT disorders.
 PS Claim 12; Page 109; 288pp; English.
 XX The invention relates to human polynucleotides and the polypeptides they
 CC encode. The polynucleotides can be used to express recombinant proteins
 CC for analysis, characterisation or therapeutic use, as markers for tissues
 CC in which the corresponding protein is expressed, as molecular weight
 CC markers on Southern gels, as chromosome markers or tags to identify
 CC chromosomes or to map related gene positions, to compare with endogenous
 CC DNA sequences in patients to identify potential genetic disorders, as
 CC probes to hybridise and discover novel related DNA sequences, as a source
 CC of information to derive PCR primers for genetic fingerprinting, to raise
 CC anti-protein antibodies and in gene therapy. The proteins can be used to
 CC raise antibodies or to elicit another immune response, as reagents in
 CC assays designed to quantitatively determine levels of the protein in
 CC biological fluids, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and to treat autoimmune disorders
 CC (e.g. multiple sclerosis, systemic lupus erythematosus, insulin dependent
 CC diabetes mellitus or graft-versus-host disease), anaemias, periodontal
 CC disorders (e.g. Alzheimer's disease or Parkinson's disease) and cancers.
 CC The proteins and polynucleotides are also useful as nutritional sources
 CC or supplements, e.g. as carbon, nitrogen or carbohydrate sources. This
 CC sequence represents a human polynucleotide of the invention.
 XX Sequence 1238 BP; 387 A; 336 C; 266 G; 249 T; 0 U; 0 Other;
 SQ
 Query Match 45.4%; Score 1011.2; DB 9; Length 1238;
 Best Local Similarity 97.9%; Pred. No. 1.3e-225;
 Matches 1226; Conservative 0; Mismatches 3; Indels 23; Gaps 19;
 QY 978 GAGCAGCTCATCAACCCCTTTGAGAGAGATGATGATTTTGGAGACCACTGGAATTGTC 1037

Db 62 TCTCTCCTCTCTCTGTGCTGCGAGGACACATCTCAAGCTGCTGTATGAGAAATTCCTT 122

Qy 216 ATCTTCCTGCTGCTACTACTACATCAATCCGCTTTATTTAAGGCTGGCCCTACAGAAAGA 275

Db 122 GTCCTCATTTCCCTACTACTATTCANTCGCTGGACCTCTACAAATGATTTCTTCAAGTAT 181

Qy 276 CAACACCTGATGTTTGAAGAAACTGACTCTGTATTTGCGACAGCTACATCCAGCTATCCC 335

Db 182 CAGACACTTGTGTTGAGAACTGCTGTGATCTGCACTCGCACACTTCATCACTCATCTCCT 241

Qy 336 ATTTCTCTTGCTGGCTGGCTTTCTACCTGACGCTGTGTCGACCCGCTGTGAAACCACTAC 395

Db 242 AATATCCTTGCTGTGGGTTTCTAATGTATCAATTGGTGTGAGCCCTGTGTGAGCACTAC 301

Qy 396 GAGAACCTGGCCGTGGCCCGACCCGCTCATGAGCCTGTGTGCGGCTTTCGTGAAAGCAAG 455

Db 302 GAGAACCTGGCCGTGGCCCGACCCGCTCATGATCCAGAGTGTCTAGCTTCGTGAGGGCAAG 361

Qy 456 GACGAGCAAGGCCGGCTGCTCGGCGGACGCTCATCCGCTAACGCAACCTGGGCAACGTG 515

Db 362 GATGAGAAAGCCGGTTGGCTGGCGGACGCTCATCCGCTAACGCAACCTGGGCAACGTG 421

Qy 516 CTCATCCGCGGACGCTCAGACCGCAGCTTACAAAGGCTTCCCAAGCCGCGACGACTG 575

Db 422 CTCATCCGCGGACGCTCAGACCGCTCGGTCTACAAAGGCTTCCCACTTTCACACCTG 481

Qy 576 GTGCAAGAGGCTTTATGACTCCGCGGCAAGACAAAGAGTTGGAAGAACTGAGCTTACA 635

Db 482 GTGCTACAGAGTTTATGACCCATGGGAAACATAGCAGTTGAGAAAGTTGGGCTTACA 541

Qy 636 CACAACATGTTCTGGGTGCTGGTGTGTTGTTGCCAATCTGTCAATGAAAGCGTGGCTT 695

Db 542 CACAACATGTTCTGGGTGCTGGTGTGTTGTTGCCAATCTGTCAATGAAAGCGCTTACTT 601

Qy 696 GAAAGTGCATCCGGGACCTTATCTGCTGCAGAGCTGCTGAAAGATGAACCTTG 755

Db 602 GAAAGTGCATCCGGGACCCGCTGCTGCAGAGCTGCTGAAATGAAGGTGTGATCTTTG 661

Qy 756 CGATCTCAGTGTGGAACAACCTGTATGCTTACGACTGATTAAGTATCCACTGGTGTATCA 815

Db 662 CGATCTCAGTGTGGAACGCTGTATGCTTACGACTGATTAAGTATCCACTGGTGTATCA 721

Qy 816 CAGTGTGTGCTGTGCGGCTGTACAGCTTCTTCGATCTGTGTCTGTGGCGGCAAGTT 875

Db 722 CAGTGTGTGCTGTGCGGCTGTATACAGCTTCTTCGATCTGTGTATCGGAGGCAAGTT 781

Qy 876 CTGAACCCGACGAGGCTTACCTGTGGCAATGACTGAGATGATCTGGTTGTGCCGTCTTACA 935

Db 782 CTGAACCCGACGAGGCTTACGAGGCTGATGATGATCTGGTTGTGCCGTCTTACA 841

Qy 936 TTCTGCAAGTTCTTCTTATGTTGGCTGGCTGAAAGGTGTGAGAGCACTCATCAACCC 995

Db 842 ATCTGCAATTTCTTATTTCTATCAATGGGCTGGCTGAAAGGTGTGAGAGCACTCATCAACCC 901

Qy 996 TTTGAGAGGATGATGATTTTGAACCAACTGATTTGTGCAAGAGAAATTTTGAAGGTG 1055

Db 902 TTTGAGAGGATGATGATTTTGAACCAACTGATTTGTGCAAGAGAAATTTTGAAGGTG 961

Qy 1056 TCCCTGTGGCTGTGAGATGAGATGACACAGGACTGTGCTGTGATGAGCCGACATGTAC 1115

Db 962 TCCCTGTGTGCTGTGAGATGAGATGACACAGAACTTGTGCTGCCATGGAACGTGACATGTAC 1021

Qy 1116 TGAATTAAGCCGAGCAACAGCCGCCCTTACAGAGCTCTTCTGCAAGTCTGTGCGGCAAT 1175

Db 1022 TGAAGAGGAGGAGGCTCAGCGGCTCAGCGGCTTACAGAGCTCTTCTGCAAGTCTGTGCGGCAAT 1081

Qy 1176 TCCCTTAATGGGCTCAGACTTCAACATATAGCTTGAACAAAGAGATGAGATTTCAGGCC 1235

Db 1082 TCCCTTAATGGGCTCAGACTTCAACATATAGCTTGAAGAAAGAGACTTGAAGCTTTGGTCA 1141

Qy 1236 AATCAGAGGACGAGAGATGCTACGCTG-----GATCATTTGGCCGCTTC 1283

Db	1142	AAAGAGAGGCTGACACGGATPAGAAAGAGTGCTATTGACGACACCTAATGGCTGCTTC	1201
QY	1284	CTAGGCGCTCAGTCCCATGATCACCATTCTCCAGGGCAAACTCAAGGACCAAACTACTG	1343
Db	1202	TTTGGAGCTCAACCCAAAACTACATCTTCTCCTTGAAAGCTTTAAAGCCAAACTATTG	1261
QY	1344	TGCGCCAAAGAGGAATCCCTTCTCCACAGAGGCGCTGCCAAAACCAAGGACCCAA	1403
Db	1262	TGTTCTTAAAG-----ACCCCTCTCTGAAAGGCAAGTGAAGATGCCAAACGAAAAAC	1315
QY	1404	CAGAACGTTAGGGGCGAAGAAAGACAAACAAAGCGCTGAGACTTTAAGCTGTGGACGCGCTC	1463
Db	1316	CAGAAAG-----ATGCTGGAATTTAAAGGCTGGACTTCTTG	1354
QY	1464	AAGTCGTGGCCCACTGTATCAGAGGCCAGGTAATCAAGTGGCCCAACAGCGCCCTTAGC	1523
Db	1355	AAATGTGTTCCAAAGTTTAAAGAGAGGCTCCCATTTGGGCCCAAGGCAACCAAGAGC	1414
QY	1524	CCCACTTC	1531
Db	1415	CACCTTAC	1422

RESULT 9	AAZ21226	standard; DNA; 16125 BP.
ID	AAZ21226	
XX		
AC	AAZ21226;	
XX		
DT	22-NOV-1999	(first entry)
XX		
DE	Human CG1CE genomic DNA sequence.	
XX		
XX	CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;	
KW	age-related macular dystrophy; ss.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	
FT	intron	Location/Qualifiers
FT		1..1513
FT		/*tag= a
FT		/number=1
FT	exon	1514..1570
FT		/*tag= b
FT		/number=1
FT		1571..3278
FT	intron	/*tag= c
FT		/number=2
FT		3279..3466
FT		/*tag= d
FT		/number=2
FT	exon	3315..15787
FT		/*tag= e
FT		/product= "CG1CE protein"
FT		/note= "contains introns"
FT	CDS	3467..6616
FT		/*tag= f
FT		/number=3
FT	exon	6617..6711
FT		/*tag= g
FT		/number=3
FT	intron	6712..7227
FT		/*tag= h
FT		/number=4
FT	exon	7228..7461
FT		/*tag= i
FT		/number=4
FT	intron	7462..8353
FT		/*tag= j
FT		/number=5
FT	exon	8354..8508
FT		/*tag= k
FT		/number=5

FT	Intron		8509.	.8896	/tag= l /number= 6	
FT	exon		8897.	.8974	/tag= m /number= 6	
PT	Intron		8975.	.9676	*tag= n /number= 7	
FT	exon		9677.	.10032	*tag= o /number= 7	
PT	Intron		10033.	.11212	/ <tag>= p /number= 8</tag>	
FT	exon		11213.	.11293	*tag= q /number= 8	
PT	Intron		11294.	.11556	*tag= r /number= 9	
FT	exon		11557.	.11708	*tag= s /number= 10	
PT	Intron		11709.	.13918	*tag= t /number= 11	
FT	exon		13919.	.14557	*tag= u /number= 10	
PT	Intron		14558.	.15768	*tag= v /number= 11	
PT	intron		14558.	.15768	*tag= w /number= 11	
FN	WO9943695-A1.					
PD	02-SEP-1999.					
PB	22-FEB-1999;	99MO-US003790.				
PR	25-FEB-1998;	98US-.0075941P.				
PA	(MERI) MERCK & CO INC. (UYUP-) UNIV UPSPALA.					
PI	Petrukhin K,	Gaekey CT, Metzker M, Wadelius C; WPI; 1999-540560/45.				
DR	P-PBDJ; AAV29953.					
XX	Human and mouse polynucleotides encoding CGICE polypeptides.					
PS	Claim 2; Fig 1; 67pp; English.					
CC	The present sequence represents the human CGICE gene, which when mutated is responsible for Beal's macular dystrophy (BMD). Polynucleotides encoding CGICE are useful for diagnosing whether a patient carries a mutation in the CGICE gene. Normal and mutated CGICE proteins are useful for identifying activators and/or inhibitors of these proteins, in order to treat BMD. The CGICE gene offers a simpler and cheaper method of diagnosing BMD without the need for the presence of the patient. The gene may also be useful to discovering the genetic cause of age-related macular dystrophy					
SQ	Sequence 16125 BP; 3988 A; 4175 G; 4215 T; 3726 U; 21 Other;					
Query Match	28 %;	Score 642.2;	DB 2:	Length 16125;		
Best Local Similarity	99.5%;	Folded No. 4.7e-139;				
Mismatch	644;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0	

QY	1202	CAGCCTGAACAAAGAGGAGATGAGATTCCAGGCCCATCAGGAGGACGAGAGAGATGCTCA	1261
Db	13916	CAGCCTGAACAAAGAGGAGATGAGATTCCAGGCCCATCAGGAGGACGAGAGAGATGCTCA	13975
QY	1262	CGCTGGCATCATTTGGCCGCTTCTCTAAGGCGTTGCGAGTCCCATGTATCACCATCTTCCAGGGC	1321
Db	13976	CGCTGGCATCATTTGGCCGCTTCTCTAAGGCGTTGCGAGTCCCATGTATCACCATCTTCCAGGGC	14035
QY	1322	AAACTCAAGACCAAACTATTGTGGCCCAAGAGGAAATCCCTTCTCACAAGAGGCTGCC	1381
Db	14036	AAACTCAAGACCAAACTATTGTGGCCCAAGAGGAAATCCCTTCTCACAAGAGGCTGCC	14095
QY	1382	CAAAAACCAAGGCAAGCCCAAGAACGTTAAGGGCCAGAAAGACAAACAGGCTGAA	1441
Db	14096	CAAAAACCAAGGCAAGCCCAAGAACGTTAAGGGCCAGAAAGACAAACAGGCTGAA	14155
QY	1442	GCTTAAAGCTGTGGACGCGCTTCAAGTCTGGCCCATGTATCAGAGGCCAGGCTACTACAG	1501
Db	14156	GCTTAAAGCTGTGGACGCGCTTCAAGTCTGGCCCATGTATCAGAGGCCAGGCTACTACAG	14215
QY	1502	TGCCCCCAAGACCGCCCTCAGCCCCCACTCCCATGTTCTTCCCTTGAACCATCAGCGCC	1561
Db	14216	TGCCCCCAAGACCGCCCTCAGCCCCCACTCCCATGTTCTTCCCTTGAACCATCAGCGCC	14275
QY	1562	GTCAAACTCTCACAAGTGTACAGAGCATAGACCAACAAAGCAAAAGCTTAAAGACTGTGAG	1621
Db	14276	GTCAAACTCTCACAAGTGTACAGAGCATAGACCAACAAAGCAAAAGCTTAAAGACTGTGAG	14335
QY	1622	TTCTGGGGCCCAAGAAAAGTTTGAATTGCTCTCAGAGACGATGGGGCCTTGATGGAGCA	1681
Db	14336	TTCTGGGGCCCAAGAAAAGTTTGAATTGCTCTCAGAGACGATGGGGCCTTGATGGAGCA	14395
QY	1682	CCCGAAGATTTCTCAAGTAGAGGAAAACTGTGAGATTAACTGTACGGAATATGCAAA	1741
Db	14396	CCCGAAGATTTCTCAAGTAGAGGAAAACTGTGAGATTAACTGTACGGAATATGCAAA	14455
QY	1742	GATCCCGGAAATTCACCTCAAGAAACCTTTGGAACATGCACCAACCAATACACACTAC	1801
Db	14456	GATCCCGGAAATTCACCTCAAGAAACCTTTGGAACATGCACCAACCAATACACACTAC	14515
QY	1802	ACTCAAAAGATCACATGAGATCTTATTGGGCGCTTGAAAAACAGGAGATG	1848
Db	14516	ACTCAAAAGATCACATGAGATCTTATTGGGCGCTTGAAAAACAGGAGATG	14562
RESULT 10			
ABAL4559	ID	ABAL4559 standard; DNA; 16650 BP.	
XX	AC	ABAL4559;	
XX	DT	23-JAN-2002 (first entry)	
XX	DE	Human nervous system related polynucleotide SEQ ID NO 6890.	
XX	XX	Human; nootropic; neuroprotective; cytosolic; dermatological; virologic;	
KM	KM	immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;	
KM	KM	antiiparkinsonian; antistickling; antihaemic; antirheumatic; cancer;	
KM	KM	antitubercular; hepatotropic; cerebroprotective; antiinflammatory;	
KM	KM	antitubercular; antidiabetic; antileper; anticonvulsant; antifungal;	
KM	KM	antiparasitic; cardiac; immune disorder; cardiovascular disorder;	
KM	KM	neurological disease; infection; nephrotropic; gene therapy; vaccine; de.	
OS	OS	Homo sapiens.	
XX	XX	WO200159063-A2.	
XX	PD	16-AUG-2001.	
XX	PE	17-JAN-2001; 2001MO-US001334.	
XX	FX	31-JAN-2000; 2000US-0179065P.	

PR	04-FEB-2000	2000US-0180628P
PR	24-FEB-2000	2000US-0184664P
PR	02-MAR-2000	2000US-0186350P
PR	16-MAR-2000	2000US-0189874P
PR	17-MAR-2000	2000US-0190076P
PR	18-MAR-2000	2000US-0198123P
PR	19-MAR-2000	2000US-0205515P
PR	07-JUN-2000	2000US-0209567P
PR	28-JUN-2000	2000US-0214886P
PR	30-JUN-2000	2000US-0215135P
PR	07-JUL-2000	2000US-0216580P
PR	07-JUL-2000	2000US-0216880P
PR	11-JUL-2000	2000US-0217487P
PR	14-JUL-2000	2000US-0219466P
PR	26-JUL-2000	2000US-0214290P
PR	26-JUL-2000	2000US-0220963P
PR	14-AUG-2000	2000US-0220964P
PR	14-AUG-2000	2000US-0224518P
PR	14-AUG-2000	2000US-0228213P
PR	14-AUG-2000	2000US-0228211P
PR	14-AUG-2000	2000US-0225266P
PR	14-AUG-2000	2000US-0225267P
PR	14-AUG-2000	2000US-0225268P
PR	14-AUG-2000	2000US-0225470P
PR	14-AUG-2000	2000US-0225470P
PR	14-AUG-2000	2000US-0225757P
PR	14-AUG-2000	2000US-0225758P
PR	14-AUG-2000	2000US-0225759P
PR	16-AUG-2000	2000US-0226279P
PR	22-AUG-2000	2000US-0226681P
PR	22-AUG-2000	2000US-0226868P
PR	22-AUG-2000	2000US-0227182P
PR	23-AUG-2000	2000US-0227009P
PR	30-AUG-2000	2000US-0228924P
PR	01-SEP-2000	2000US-0229287P
PR	01-SEP-2000	2000US-0229343P
PR	01-SEP-2000	2000US-0229343P
PR	05-SEP-2000	2000US-0229345P
PR	05-SEP-2000	2000US-0229509P
PR	06-SEP-2000	2000US-0229513P
PR	06-SEP-2000	2000US-0230437P
PR	08-SEP-2000	2000US-0230438P
PR	08-SEP-2000	2000US-0231242P
PR	08-SEP-2000	2000US-0231423P
PR	08-SEP-2000	2000US-0231413P
PR	08-SEP-2000	2000US-0231414P
PR	08-SEP-2000	2000US-0232080P
PR	08-SEP-2000	2000US-0232081P
PR	14-SEP-2000	2000US-0231968P
PR	14-SEP-2000	2000US-0232397P
PR	14-SEP-2000	2000US-0232398P
PR	14-SEP-2000	2000US-0232398P
PR	14-SEP-2000	2000US-0232400P
PR	14-SEP-2000	2000US-0232401P
PR	14-SEP-2000	2000US-0232401P
PR	14-SEP-2000	2000US-0233063P
PR	14-SEP-2000	2000US-0233064P
PR	21-SEP-2000	2000US-0233065P
PR	21-SEP-2000	2000US-0234223P
PR	21-SEP-2000	2000US-0234274P
PR	25-SEP-2000	2000US-0234997P
PR	25-SEP-2000	2000US-0234998P
PR	26-SEP-2000	2000US-0234998P
PR	27-SEP-2000	2000US-0235484P
PR	27-SEP-2000	2000US-0235833P
PR	29-SEP-2000	2000US-0235836P
PR	29-SEP-2000	2000US-0236327P
PR	29-SEP-2000	2000US-0236367P
PR	29-SEP-2000	2000US-0236368P
PR	29-SEP-2000	2000US-0236369P
PR	02-OCT-2000	2000US-0236370P
PR	02-OCT-2000	2000US-0236802P
PR	02-OCT-2000	2000US-0237037P

PR	02-OCT-2000	2000US-0237033P.
PR	02-OCT-2000	2000US-0237039P.
PR	02-OCT-2000	2000US-0237040P.
PR	13-OCT-2000	2000US-0239353P.
PR	13-OCT-2000	2000US-0239373P.
PR	20-OCT-2000	2000US-0240566P.
PR	20-OCT-2000	2000US-0241785P.
PR	20-OCT-2000	2000US-0241785P.
PR	20-OCT-2000	2000US-0241787P.
PR	20-OCT-2000	2000US-0241809P.
PR	20-OCT-2000	2000US-0241826P.
PR	20-OCT-2000	2000US-0242221P.
PR	01-NOV-2000	2000US-0244617P.
PR	08-NOV-2000	2000US-0246474P.
PR	08-NOV-2000	2000US-0246475P.
PR	08-NOV-2000	2000US-0246476P.
PR	08-NOV-2000	2000US-0246477P.
PR	08-NOV-2000	2000US-0246478P.
PR	08-NOV-2000	2000US-0246523P.
PR	08-NOV-2000	2000US-0246524P.
PR	08-NOV-2000	2000US-0246525P.
PR	08-NOV-2000	2000US-0246526P.
PR	08-NOV-2000	2000US-0246537P.
PR	08-NOV-2000	2000US-0246538P.
PR	08-NOV-2000	2000US-0246532P.
PR	08-NOV-2000	2000US-0246609P.
PR	08-NOV-2000	2000US-0246610P.
PR	08-NOV-2000	2000US-0246611P.
PR	17-NOV-2000	2000US-0246613P.
PR	17-NOV-2000	2000US-0249207P.
PR	17-NOV-2000	2000US-0249208P.
PR	17-NOV-2000	2000US-0249209P.
PR	17-NOV-2000	2000US-0249210P.
PR	17-NOV-2000	2000US-0249211P.
PR	17-NOV-2000	2000US-0249212P.
PR	17-NOV-2000	2000US-0249213P.
PR	17-NOV-2000	2000US-0249214P.
PR	17-NOV-2000	2000US-0249215P.
PR	17-NOV-2000	2000US-0249216P.
PR	17-NOV-2000	2000US-0249217P.
PR	17-NOV-2000	2000US-0249241P.
PR	17-NOV-2000	2000US-0249244P.
PR	17-NOV-2000	2000US-0249245P.
PR	17-NOV-2000	2000US-0249264P.
PR	17-NOV-2000	2000US-0249265P.
PR	17-NOV-2000	2000US-0249297P.
PR	17-NOV-2000	2000US-0249299P.
PR	17-NOV-2000	2000US-0249300P.
PR	01-DEC-2000	2000US-0250391P.
PR	01-DEC-2000	2000US-0251160P.
PR	05-DEC-2000	2000US-0251030P.
PR	05-DEC-2000	2000US-0251988P.
PR	06-DEC-2000	2000US-0256719P.
PR	08-DEC-2000	2000US-0251479P.
PR	08-DEC-2000	2000US-0251856P.
PR	08-DEC-2000	2000US-0251868P.
PR	08-DEC-2000	2000US-0251989P.
PR	08-DEC-2000	2000US-0251989P.
PR	11-DEC-2000	2000US-0251990P.
PR	05-JAN-2001	2001US-0259678P.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
XX	Rosen CA, Barash SC, Ruben SM;	
XX	WPI, 2001-541565/60.	
DR		
XX	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
XX	useful for preventing, diagnosing and/or treating nervous system cancers	
XX	and metastases.	
XX		

PS Disclosure; SEQ ID NO 6890; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-AB21534) and proteins
CC (AB14678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 16650 BP; 4168 A; 4271 C; 4308 G; 3903 T; 0 U; 0 Other;

Query Match 28.8%; Score 642.2; DB 5; Length 16650;
Best Local Similarity 99.5%; Pred. No. 4.7e-139;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1202 CAGCTGAACAAGAGGAGTTCAGCCCAATCAGAGAGAGGAGATGCTCA 1261
DB 11895 CAGCTGAACAAGAGGAGTTCAGCCCAATCAGAGAGAGGAGATGCTCA 11954
QY 1262 CGCTGCAATCTGGCCGCTTCTAGGCTGAGTCCCATGTCACATCTCCAGAGGC 1321
DB 11955 CGCTGCAATCTGGCCGCTTCTAGGCTGAGTCCCATGTCACATCTCCAGAGGC 12014
QY 1222 AAATCAAGAACAATCTAGTGTGGCCCAAGAGGAAATCCCTTCTCAAGAGGCTGCC 1381
DB 12015 AAATCAAGAACAATCTAGTGTGGCCCAAGAGGAAATCCCTTCTCAAGAGGCTGCC 12074
QY 1282 CAAAACCAAGAGGAGGCAACGAAGTTAGGGCCAGAGAGCAACAAGGCTGGAA 1441
DB 12075 CAAAACCAAGAGGAGGCAACGAAGTTAGGGCCAGAGAGCAACAAGGCTGGAA 12134
QY 1442 GCTTAAGGCTGTGAGCGCTTCAAGTGTGCGCCCACTGATCAGAGGCGAGCTACTACAG 1501
DB 12135 GCTTAAGGCTGTGAGCGCTTCAAGTGTGCGCCCACTGATCAGAGGCGAGCTACTACAG 12194
QY 1502 TGCCCAAGAGCGCCCTCAGCGCCCACTCCAGATTTCTTCCCTAGAACATCAGCGCC 1561
DB 12195 TGCCCAAGAGCGCCCTCAGCGCCCACTCCAGATTTCTTCCCTAGAACATCAGCGCC 12254
QY 1562 GTCAAGCTTCAAGTGTGACAGGATGACACCAAAAGCTTAAAGACTGTGAG 1621
DB 12255 GTCAAGCTTCAAGTGTGACAGGATGACACCAAAAGCTTAAAGACTGTGAG 12314
QY 1622 TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGATGAGCA 1681
DB 12315 TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGATGAGCA 12374
QY 1682 CCCAGAAAGTATCTCAAGTGAAGAGAAAGTGTGAGTTTAACTGACGATATGCCAGA 1741
DB 12375 CCCAGAAAGTATCTCAAGTGAAGAGAAAGTGTGAGTTTAACTGACGATATGCCAGA 12434
QY 1742 GATCCCGCAAAATCACTCAAGAACTTTTGAACATACCAACCAATACACACTAC 1801
DB 12435 GATCCCGCAAAATCACTCAAGAACTTTTGAACATACCAACCAATACACACTAC 12494
QY 1802 ACTCAAGATCAATGATCTTATTGGGCTTGAGAAACAGGATG 1848
DB 12495 ACTCAAGATCAATGATCTTATTGGGCTTGAGAAACAGGATG 12541

RESULT 11
AB14557

ID ID ABA14557 standard; DNA; 18530 BP.
XX
XX ABA14557;
AC
XX 23-JAN-2002 (first entry)
DT
XX
XX Human nervous system related polynucleotide SEQ ID NO 6888.
DE
XX
XX Human; noctropic; neuroprotective; cytostatic; dermatological; virocid;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antispasmodic; antidiabetic; antianemic; antirheumatic;
KW antitubercular; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antidiabetic; antidiabetic; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001334.
PF
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.

Db 12195 TGCCCCACAGAGCCCTCCAGCCCCACTCCCATGTTCTTCCCCCTAGAACCATCAGCGCC 12254
Qy 1562 GTCAAGCTTCAGAGTGCATGACAGCATAGACACCAAGAAAGCTTAAGACTGTAG 1621
Db 12255 GTCAAGCTTCAGAGTGCATGACAGCATAGACACCAAGAAAGCTTAAGACTGTAG 12314
Qy 1622 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCATGGGGCTTGATGAGCA 1681
Db 12315 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCATGGGGCTTGATGAGCA 12374
Qy 1682 CCCAAGATATCTCAAGTAGAGAGAAAACTGTGAGTTTAACCTGACGATATGCCAGA 1741
Db 12375 CCCAAGATATCTCAAGTAGAGAGAAAACTGTGAGTTTAACCTGACGATATGCCAGA 12434
Qy 1742 GATCCCGGAAAAATACCTCTCAAGAAACCTTTGGAAACATCACAACCAATACACTAC 1801
Db 12435 GATCCCGGAAAAATACCTCTCAAGAAACCTTTGGAAACATCACAACCAATACACTAC 12494
Qy 1802 ACTCAAGATCAGATGATCCTTATTTGGGCTTGAAAAACAGGATG 1848
Db 12495 ACTCAAGATCAGATGATCCTTATTTGGGCTTGAAAAACAGGCTGTG 12541

RESULT 12
AB14558
ID AB14558 standard; DNA; 18537 BP.
XX
AC AB14558;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6889.
XX
KW Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; anticholinerg; antianaemic; antirheumatic; cancer;
KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; neurotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
PD 16-AUG-2001.
PF 17-JAN-2001; 2001WO-US001334.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218230P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234987P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236357P.
PR 29-SEP-2000; 2000US-0236358P.
PR 29-SEP-2000; 2000US-0236359P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242211P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.

PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250391P.
PR	01-DEC-2000;	2000US-0251160P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	06-DEC-2000;	2000US-0256719P.
PR	08-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251999P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0259678P.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Barash SC, Ruben SM;	
PI	WPI; 2001-541565/60.	
DR	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating nervous system cancers	
PT	and metastases.	
PS	Disclosure; SEQ ID NO 6889; 1701bp + Sequence Listing; English.	
XX		
CC	The invention relates to novel genes (ABAI1004-ABA21534) and proteins	
CC	(ABAI1678-ABAI8001) useful for preventing, treating or ameliorating	
CC	medical conditions e.g. by protein or gene therapy. The genes are	
CC	isolated from a range of human tissues disclosed in the specification.	
CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful in	
CC	the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and	
CC	ovarian cancer and other cancers of the adrenal gland, bone, bone marrow	
CC	breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune	
CC	disorders e.g. Addison's disease, allergies, autoimmune haemolytic	
CC	anaemia, autoimmune rheumatoid arthritis and ulcerative colitis; (c)	
CC	multiple sclerosis, thrombotic thrombocytopenic purpura; (d) neurologic disorders such as myocardial ischaemias; (e) wound healing	
CC	cardiovascular disorders such as myocardial ischaemias; (f) neurological diseases e.g. cerebral anoxia and epilepsy; and (g)	
CC	(e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)	
CC	infections diseases such as viral, bacterial, fungal and parasitic	
CC	infections. Note: The sequence data for this patent did not form part of	
CC	the printed specification, but was obtained in electronic format directly	
XX	from WIPO at ftp.wipo.int/pub/published_pct_sequences	
SQ	Sequence 18537 BP; 4519 A; 4800 C; 4983 G; 4235 T; 0 U; 0 Other;	
Query Match	28.7%; Score 640.6; DB 5; Length 18537;	

[illegible]

XX (HUMA-) HUMAN GENOME SCI INC.
 PA Birtse CE, Rosen CA;
 PI WPI; 2002-122018/16.
 DR P-PSDB; ABB89288.
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders.
 PS Claim 4; SEQ ID NO 259; 2081bp + Sequence Listing; English.
 XX The invention relates to novel genes (ABL9449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 1717 BP; 464 A; 460 C; 417 G; 367 T; 0 U; 9 Other;

Query Match 26.4%; Score 587.8; DB 6; Length 1717;
 Best Local Similarity 98.8%; Pred. No. 9.8e-127;
 Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1242 GAGACGAGGAGGATGCTCAGGCTGGGCTTCTTGGCCGCTTCCATGAGGCTGATCCCAT 1301
 DB 1 GAGACGAGGAGGATGCTCAGGCTGGGCTTCTTGGCCGCTTCCATGAGGCTGATCCCAT 60
 QY 1302 GATCACCATCTCTCCAGGAGCAACTCAAGACCAAACTACTGTGTGGCCCAAGAGGAATCC 1361
 DB 61 GATCACCATCTCTCCAGGAGCAACTCAAGACCAAACTACTGTGTGGCCCAAGAGGAATCC 120
 QY 1362 CTCTCCACGAGGCTGCTCCCAAAACCAAGGAGGAGCAAGAGCTTGTGGGCGCAG 1421
 DB 121 CTCTCCACGAGGCTGCTCCCAAAACCAAGGAGGAGCAAGAGCTTGTGGGCGCAG 179
 QY 1422 GAAGACCAAGAGGCTGAGAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 1481
 DB 180 GAAGACCAAGAGGCTGAGAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 239
 QY 1482 CAGAGGCGAGGCTACTCAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1541
 DB 240 CAGAGGCGAGGCTACTCAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 299
 QY 1542 CCCCTAGAACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1601
 DB 300 CCCCTAGAACCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 359
 QY 1602 AAAAGCTTAAAGAGCTGTGAGTTCGTGGGGCCCAAGAAAGTTTGAATTCCTTCAGAGAGC 1661
 DB 360 AAAAGCTTAAAGAGCTGTGAGTTCGTGGGGCCCAAGAAAGTTTGAATTCCTTCAGAGAGC 419
 QY 1662 GATGGGGGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1721
 DB 420 GATGGGGGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 479
 QY 1722 AACCTGACGATATGCGAGAGATCCCGAAATATCACTCAAGAAAGACCTTTTGAACATCA 1781

DB 480 AACCTGACGATATGCGAGAGATCCCGAAATATCACTCAAGAAAGCTTTGGAACATCA 539
 QY 1782 CCACCAACATACACACTACACTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1841
 DB 540 CCACCAACATACACACTACACTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
 QY 1842 AGGAGATG 1848
 DB 600 AGGTCTG 606

RESULT 14
 AB061177
 ID AB061177 standard; cDNA; 2112 BP.
 XX
 AC AB061177;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE FLJ20132 fls clone COL06441 encoding sequence.

Neuroprotective; immunomodulator; cancer; chromosome 19; cytostatic;
 anti-inflammatory; gene therapy; nutritional supplement; wound; burn;
 ulcer; Alzheimer's disease; Huntington's disease;
 amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
 vulnery; gene; ss.

Homo sapiens.
 MO200231111-A2.
 18-APR-2002.
 11-OCT-2001; 2001MO-US027760.

12-OCT-2000; 2000US-00687527.
 (HYSE-) HYSEQ INC.

Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 Xue AO, Yang Y, Wehrman T, Drmanac RT;
 WPI; 2002-426278/45.
 N-PSDB; ABP43933.

New polypeptides and their encoded proteins, useful as nutritional
 PT sources or supplements, or in gene therapy, particularly for treating
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
 PT inflammation.

Claim 1; SEQ ID # 390; 357bp + Sequence Listing; English.

The invention relates to 446 newly isolated polynucleotide sequences. The
 CC activity of polynucleotides of the invention may be described as,
 CC vulnerability, neuroprotective, immunomodulator, cytostatic and anti-
 CC inflammatory. Compositions comprising nucleic acids of the invention are
 CC useful for treating a mammalian subject, or as nutritional sources or
 CC supplements. These are useful in gene therapy, particularly for treating
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
 CC inflammation. The nucleic acids and polypeptides are also useful in
 CC diagnostic and research methods. The sequences given in records AB060788-
 CC AB061233 represent polynucleotides of the invention. NOTE: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

Query Match 25.9%; Score 577.2; DB 6; Length 2112;
 Best Local Similarity 69.9%; Pred. No. 3.2e-124;
 Matches 793; Conservative 0; Mismatches 338; Indels 3; Gaps 1;

QY 103 CCAATGACATCACTTACACAAAGCAAGTGGTATATGCCGCTTAGGCTCTTCCTCCGCC 162
 Db 190 CGATGACCGCTCACTTACACAGCCGAGTGGCGAAGCCGCCCTTCGCTGGCTTCCAGC 249
 QY 163 TGGTGTGCTGGTGGGGGACAGATCTTACAGCTGCTATATGGCGAGTTCTTAATCTTC 222
 Db 250 TGGTGTGCTGGTGGGGGACAGATCTTACAGCTGCTATATGGCGAGTTCTTCCTCC 309
 QY 223 TGGTGTGCTGGTGGGGGACAGATCTTACAGCTGCTATATGGCGAGTTCTTCCTCC 282
 Db 310 TGGGTTCTACATGGCGCTGAGTGGCTTACCGCTTGTGCTACCGAAGGCGAAGC 369
 QY 283 TGAATTTGAGAAATGACTCTGATATGGCGACAGCTATCCAGCTATCCCATTTCTCT 342
 Db 370 GCTACTGAGAAAGCTTGTGATTTATGTTGACAGTATGACAGCTATCCCTGTCTCT 429
 QY 343 TGGTGTGCTGGTGGGGGACAGATCTTACAGCTGCTATATGGCGAGTTCTTCCTCC 402
 Db 430 TGGTGTGCTGGTGGGGGACAGATCTTACAGCTGCTATATGGCGAGTTCTTCCTCC 489
 QY 403 TGGTGTGCTGGTGGGGGACAGATCTTACAGCTGCTATATGGCGAGTTCTTCCTCC 462
 Db 490 TGGGTTCTACATGGCGCTGAGTGGCTTACCGCTTGTGCTACCGAAGGCGAAGC 549
 QY 463 AAGCCGCGCTGGTGGGGGACAGATCTTACAGCTGCTATATGGCGAGTTCTTCCTCC 522
 Db 550 GCGGCGCGCTTACAGCGCGCTGATGCTATGCTACAGCGCGCTTCTCGCGCTCTATCC 609
 QY 523 TGGGCGCGCTTACAGCGCGCTGATGCTATGCTACAGCGCGCTTCTCGCGCTCTATCC 582
 Db 610 TGGGCGCGCTTACAGCGCGCTGATGCTATGCTACAGCGCGCTTCTCGCGCTCTATCC 669
 QY 583 CAGGCTTTATGATCTCGCGCGAAGACACAGAGCTTGGAGAACTGAGCTTACACACA 642
 Db 670 CTGGGTTTATGATCTCGCGCGAAGACACAGAGCTTGGAGAACTGAGCTTACACACA 729
 QY 643 TGGTGTGCTGGTGGGGGACAGATCTTACAGCTGCTATATGGCGAGTTCTTCCTCC 702
 Db 730 AGTACTGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
 QY 703 GAATCGGCGCGCTTACAGCGCGCTGATGCTATGCTACAGCGCGCTTCTCGCGCTCT 762
 Db 790 GCAATCGCGCGAAGCGCGCTTACAGCGCGCTGATGCTATGCTGCTGCTGCTGCTGCT 849
 QY 763 AGTGTGACACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 822
 Db 850 AATGTGATGCTGCTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909
 QY 823 TGAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882
 Db 910 TGAATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 969
 QY 883 CAGCCAGCGCTTACAGCGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 942
 Db 970 CGGCTCAGCGGTTCAGAAAGCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1029
 QY 943 AGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002
 Db 1030 AGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1089
 QY 1003 AGATGTGATGATTTTGAAGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
 Db 1090 AGATGTGATGATTTTGAAGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149
 QY 1063 TGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1122
 Db 1150 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1209
 QY 1123 AGCCGAGGACAGCCCGCTTACAGCTGCTTC---GCGCCAGTTCGCTGAGGCTCTCT 1179
 Db 1210 CAGCCAGGCTGCGCGCGCTTACAGCTGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1269
 QY 1180 TTATGGCTTCAACTTCAACATGAGCTGTAAGAAAGAGATGATGATTCAGC 1233

Db 1270 TCCAGGCGCTCCACCTTATGACATCAGCTGGCCAAAGAGATGATGATTCAGC 1323

RESULT 15
 AB280973
 ID AB280973 standard; DNA; 2137 BP.
 AC XX
 AB280973;
 XX
 DT 15-OCT-2003 (first entry)
 XX
 DE Human bestrophin (vitelliform macular dystrophin 2-like protein 1) gene.
 XX
 KW Bestrophin; vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;
 KW antiobesity; antidiabetic; immunomodulator; hypotensive; cardiac;
 KW antilipemic; osteopathic; antinflammatory; cyostatic; obesity;
 KW energy homeostasis; metabolism; triglyceride; body-weight regulation;
 KW eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
 KW coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
 KW sleep apnea; chromosome 19.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 212..1741
 FT /*tag= a
 FT /product= "VMD2-like protein 1"
 XX
 PD W02003030922-A2.
 XX
 PA 17-Apr-2003.
 XX
 PR 09-OCT-2002; 2002MO-BE011321.
 XX
 PR 09-OCT-2001; 2001BP-00124059.
 XX
 PI (DEVE-) DEVELOPENT ENTWICKLUNGSRIOLOGISCHE FORSCH.
 XX
 DR Steuermangel A, Broenner G, Fritsch R, Eulenberger K, Ciosek T;
 DR WPI; 2003-393411/37.
 DR P-PSDB; ABR58056.
 XX
 PS New pharmaceutical compositions comprising a Bestrophin gene, polypeptide
 XX or nucleic acid, for treating, alleviating and/or preventing metabolic
 XX diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
 XX gallstones.
 XX
 PS Claim 2; Fig 4c; 85pp; English.
 XX
 CC This sequence represents the coding region for a member of the human
 CC bestrophin gene family designated vitelliform macular dystrophin 2 (VMD2)
 CC -like protein 1. The dystrophin gene family are involved in energy
 CC homeostasis and metabolism of triglycerides. The sequence can be used for
 CC the manufacture of an agent for detecting and/or verifying, for treating,
 CC alleviating and/or preventing disorders including metabolic diseases such
 CC as obesity and other body-weight regulation and related disorders such as
 CC eating disorder, cachexia, diabetes mellitus, hypertension, coronary
 CC heart disease, hypercholesterolemia, osteoarthritis, gallstones, cancers
 CC of the reproductive organs, and sleep apnea. The gene is found on human
 CC chromosome 19
 XX
 SQ Sequence 2137 BP; 399 A; 671 C; 622 G; 445 T; 0 U; 0 Other;
 XX
 Query Match 25.9%; Score 577.2; DB 8; Length 2137;
 Best Local Similarity 69.9%; Pred. No. 3.2e-124;
 Matches 793; Conservative 0; Mismatches 338; Indels 3; Gaps 1;
 QY 103 CCAATGACATCACTTACACAAAGCAAGTGGTATATGCCGCTTAGGCTCTTCCTCCGCC 162
 Db 210 CGATGACCGCTCACTTACACAGCCGAGTGGCGAAGCCGCCCTTCGCTGGCTTCCAGC 269

Search completed: March 25, 2005, 23:51:32
Job time : 1172.53 secs

Qy	163	TGCTGCTGCTGCTGGGGGAGAGACTTCAAGAGCTGCTATATGGGAGATTTCTTATCTTCC	222
Dp	270	TGCTGCTACTGGGGGGAGACATCTCAAACTCTGTGGCGAGAGCTGCTCTGCTTCC	329
Qy	223	TGCTCTGTACTACATCATCCGCTTATATTTATAGAGCTGGCCCTCAGGAACAAACAGC	282
Dp	330	TTGGGTTTACAATGAGCGCTGAGTGTGCTGCTTACCGTTTGTGTGACCGAAGGCGAAGC	389
Qy	283	TGATGTTTGAAGAACTGACTCTGTATTTGCGACAGCTACATCAGCTCATCTCCATTCT	342
Dp	390	GCTACTTGAGAGAGTTGTGATTTATTTATGTGACAGATATGCAAGCTCATCTCTGTCT	449
Qy	343	TGTGCTGGGCTTCTACGTGACGTGTGCTGTGACCCGCTGGTGGAAACATATCGAAGAC	402
Dp	450	TCTGCTCTTGGCTTTTATGTGACGCTGTGTGGAAACCGCTGTGTGAGCCATCTATATGCA	509
Qy	403	TGCGGTGGCCGACCGGCTCATATGAGCCGTGTGTGGGGCTTGTGTGAAAGCAAGACGAGC	462
Dp	510	TGCGGCTGCCGAGCGGCTCATATGTGCTGTGTGGGGGACCGTGTGACGAGCGACGAGCC	569
Qy	463	AAGCGCGGCTCTGGGCGGACAGCTCATCCGCTACGCGCAACTGGGCAACCTGTGCTATCC	522
Dp	570	GCGGCGGCTCTAACCGGGGACACATCATGTGCTATAGCAGAGGCTCTCGGCGGTGTCTATCC	629
Qy	523	TGCGAGCGTCAGACCGCAGTCTTACAAGCCTTCCGAGCGCCACAGCACTGTGTCAAG	582
Dp	630	TGCGCTCGTCAAGCACCGCGGTGTTCAGCGCTTCCACATATAGACCACTGTGTGAGG	689
Qy	583	CAGGCTTATGATCTCCGACAGAACACAGCAGTTGGAGAACTAGGCTTACCACACACAA	642
Dp	690	CTGGGTTTATATACCGCGACGAGGCGCAAGAAAGTTGAAACCTGAACTCATCTTACACA	749
Qy	643	TGTTCTGGGGTCCGCTGGGTGTGTGTTTGGCAACTGTCAATAGAGGCGTGGCTTGGAGATC	702
Dp	750	AGTACTGGGTGCTGCTGCTGTGTCTTCAACTTGGCGGACAGGCCCTCACGCGAGGGCC	809
Qy	703	GAATCCGGGACCTATCTCTGTCCAGAGCCTGTCTAAAGAGATGAAACACTTGTGCTACTC	762
Dp	810	GCATCCGGACAAACAGCGCCCTTAAAGTGTGTGCTGCTGAGAGAGCTGAATGTTTTGGGGCA	869
Qy	763	AGTGTGACACCTGTATGCTTACGACTGTGATTTAGTATCCACTGTGTATATACAGGTGG	822
Dp	870	AATGTGGATGCTCTTTCACATATGACTGATTTAAGGTACCCCTGTGTACACGACGGTGG	929
Qy	823	TGACTGTGGCGGTGATACAGCTTCTTCCGACTGTGCTATGTGGGGGAGAGTTTCTGAACC	882
Dp	930	TGACCATGCACTGTACAGCTACTTCTCTGGCTGTGCTATTTGTGTGCCAGTTCCTGACCC	989
Qy	883	CAGCCAAAGGCTTACCCTGGGCATGAGCTGTGACCTGTGTGTCGCGCTTTCACGTTCTGTC	942
Dp	990	CGGCTCAGGGTTACAAABAACACAGACTATAGACCTGTGTGTGTCCTTACACCTCTTGG	1049
Qy	943	AGTTCTTCTTATGTTGGCTGTGGCTGAAGGTGGCAGACAGCTCATCAACCCCTTTGGAG	1002
Dp	1050	AGTTCTTCTTCAAGCGCGGCTGTGCTCAAGGTATGTCGACACTCATCAACCCCTTGGAG	1109
Qy	1003	AGATATGATGATTTTATGAGACCACTGGAATTTGTGACAGGAATTTGAGAGTGTCCCTGT	1062
Dp	1110	AGGACGATGATGACTTTTATGAGACCACTTTCATGTGATGATGAAACTTCCAGGTGTCCATGC	1169
Qy	1063	TGGCTGTGATGATGATGACACGAGACTGTGCTGATGAGACCGGACATGTACTGAAATA	1122
Dp	1170	TGGCAGTGTGAGAGATGTATGTATGACTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	1229
Qy	1123	AGCCGAGCCACAAGCCCCCTTACACAGCTGTTC--GCGCAGTTCCGTGTGAGCTCT	1179
Dp	1230	CAGCGAGGCTGTGCGCCCATATACAGCGGCTATGTCTTCAAGCTGTGGGAGCGCTTCT	1289
Qy	1180	TTATGGGCTTCACTTCAACATCACTCAGCTGAAACAAAGAGATGAGATTTCCAGC	1233
Dp	1290	TTCAGGGCTTCACTTTGACATCAAGCTGTGGCCAAAGAAAGATGAGATTTCCAGC	1343

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2005, 07:39:17 ; Search time 1241.65 Seconds
(without alignments)
10698.156 Million cell updates/sec

Title: US-09-622-964A-2
Perfect score: 2229
Sequence: 1 caggagagtcaccacgcctca.....aaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 5552208 seqs, 2979665951 residues

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1754.8	78.7	1758	US-10-492-032-1	Sequence 1, Appl1
2	1247.2	56.0	1263	US-09-746-783-3	Sequence 3, Appl1
3	587.8	26.4	1717	US-10-264-237-859	Sequence 259, App
4	577.2	25.9	2137	US-10-492-032-3	Sequence 3, Appl1
5	576.8	25.9	1530	US-10-492-032-30	Sequence 129, Ap
6	459.2	20.6	2500	US-10-108-2604-1299	Sequence 129, Ap
7	446.2	20.0	1422	US-10-492-032-7	Sequence 7, Appl1
8	423	19.0	2028	US-10-492-032-5	Sequence 5, Appl1
9	350.4	15.7	1326	US-10-071-766-10	Sequence 10, Appl
10	215.4	9.7	1350	US-09-768-826-16	Sequence 16, Appl
11	215.4	9.7	1350	US-10-874-484-16	Sequence 16, Appl

C	12	213	9.6	620	10	US-09-814-353-19504	Sequence 19504, A
C	13	189.2	8.5	1292	14	US-10-198-846-11070	Sequence 11070, A
C	14	175	7.9	1198	9	US-09-880-107-2174	Sequence 2174, Ap
C	15	161.4	7.2	578	13	US-10-027-632-134530	Sequence 134530, A
C	16	161.4	7.2	578	13	US-10-027-632-134530	Sequence 134530, A
C	17	149.4	6.7	748	13	US-10-027-632-134529	Sequence 134529, A
C	18	149.4	6.7	748	13	US-10-027-632-134529	Sequence 134529, A
C	19	135.2	6.1	539	16	US-10-029-386-7583	Sequence 7583, Ap
C	20	134.2	6.0	235	16	US-10-029-386-2183	Sequence 2183, A
C	21	125.4	5.6	2914	18	US-10-723-860-7441	Sequence 7441, Ap
C	22	125.4	5.6	2914	18	US-10-723-860-8303	Sequence 8303, Ap
C	23	116	5.2	853	14	US-10-198-846-11346	Sequence 11346, A
C	24	115.4	5.2	1954	18	US-10-723-860-7576	Sequence 7576, A
C	25	104	4.7	462	14	US-10-198-846-2561	Sequence 2561, Ap
C	26	103.6	4.6	3186778	13	US-10-027-632-174961	Sequence 174961, A
C	27	103.6	4.6	3186778	17	US-10-027-632-174961	Sequence 174961, A
C	28	102	4.6	615	14	US-10-198-846-7843	Sequence 7843, Ap
C	29	98	4.4	507	10	US-09-764-872-195	Sequence 195, App
C	30	74.2	3.3	239	18	US-10-637-855-110	Sequence 110, App
C	31	66.2	3.0	442	10	US-09-814-353-14343	Sequence 14343, A
C	32	66	3.0	748	13	US-10-027-632-134529	Sequence 134529, A
C	33	66	3.0	748	17	US-10-027-632-134529	Sequence 134529, A
C	34	65.4	2.9	260	10	US-09-814-353-1601	Sequence 1601, Ap
C	35	65.4	2.9	260	10	US-09-814-353-13563	Sequence 13563, A
C	36	65.4	2.9	463	10	US-09-814-353-7959	Sequence 7959, Ap
C	37	64	2.9	751	14	US-10-198-846-1862	Sequence 1862, Ap
C	38	63.2	2.8	298	10	US-09-814-353-806	Sequence 806, App
C	39	63.2	2.8	298	10	US-09-814-353-7178	Sequence 7178, Ap
C	40	60	2.7	65	10	US-09-908-975-16358	Sequence 16358, A
C	41	54.4	2.4	481	18	US-09-908-975-29941	Sequence 29941, A
C	42	52.4	2.4	1274	18	US-10-767-701-19477	Sequence 19477, A
C	43	49.6	2.2	1263	10	US-10-437-963-87249	Sequence 87249, A
C	44	49.2	2.2	1326	10	US-09-746-783-3	Sequence 3, Appl1
C	45	49.2	2.2	1326	13	US-10-071-766-10	Sequence 10, Appl1

ALIGNMENTS

RESULT 1
US-10-492-032-1
Sequence 1, Application US/10492032
Publication No. US20050049212A1
GENERAL INFORMATION:
APPLICANT: Steuermann, Arnd
APPLICANT: Brommer, Gunter
APPLICANT: Fritsch, Rüdiger
APPLICANT: Eulenberg, Karsten
TITLE OF INVENTION: Cloasek, Thomas
TITLE OF INVENTION: Regulation and Bistrophin Homologous Proteins Involved in the
FILE REFERENCE: 2923-606
CURRENT APPLICATION NUMBER: US/10/492,032
PRIOR FILING DATE: 2004-04-08
PRIOR APPLICATION NUMBER: PCT/EP02/11321
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1758
TYPE: DNA
ORGANISM: Homo sapiens
US-10-492-032-1
Query Match 78.7%; Score 1754.8; DB 19; Length 1758;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1756; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 105 ATGACCATCACTTACACAGCAAGTGGCTATAGCCCGCTTCTCTCCGCTG 164
Db 1 ATGACCATCACTTACACAGCAAGTGGCTATAGCCCGCTTCTCTCCGCTG 60

QY 165 CTGCTGTGCTGGGCGGCGAGCATCTCAAGCTGTATGATGCGAGTTCTTAATCTTCCG 224
 DB 61 CTGCTGTGCTGGGCGGCGAGCATCTCAAGCTGTATGATGCGAGTTCTTAATCTTCCG 120
 QY 225 CTCTGTACTACATCATCTCCGCTTTATTTAAGGCTTGCCCTTCACGGAAGAACAGCTG 284
 DB 121 CTCTGTACTACATCATCTCCGCTTTATTTAAGGCTTGCCCTTCACGGAAGAACAGCTG 180
 QY 285 ATGTTTGAAGAACTGACTGTATGCGACAGCTCATCTCAAGCTGTATGCGAGTTCTTCC 344
 DB 181 ATGTTTGAAGAACTGACTGTATGCGACAGCTCATCTCAAGCTGTATGCGAGTTCTTCC 240
 QY 345 GTGCTGGGCTTCTCAAGTGAAGCTGTGCTGAGCCCGCTGGTGAAGCAAGTACAGAACTG 404
 DB 241 GTGCTGGGCTTCTCAAGTGAAGCTGTGCTGAGCCCGCTGGTGAAGCAAGTACAGAACTG 300
 QY 405 CCGTGCCCGACCGGCTCTATGAGCTGTGCTGGCTTCTGCGAAGGCAAGGACGAGCA 464
 DB 301 CCGTGCCCGACCGGCTCTATGAGCTGTGCTGGCTTCTGCGAAGGCAAGGACGAGCA 360
 QY 465 GGCCTGGCTGTGCGGCGGCGACGCTCATCTCCGCTAAGCGCAAGCTGGGCAAGTGTCTG 524
 DB 361 GGCCTGGCTGTGCGGCGGCGACGCTCATCTCCGCTAAGCGCAAGCTGGGCAAGTGTCTG 420
 QY 525 CGCAGCGTTCAGCAGCGAGCTTCAAGAGGCTTCCCGACGCGCCAGCAGCTGGTGCAGCA 584
 DB 421 CGCAGCGTTCAGCAGCGAGCTTCAAGAGGCTTCCCGACGCGCCAGCAGCTGGTGCAGCA 480
 QY 585 GGCCTTATGACCTCCGCGGAGAAACAAGAGTTGGAGAACTGAGGCTTCCACACACATG 644
 DB 481 GGCCTTATGACCTCCGCGGAGAAACAAGAGTTGGAGAACTGAGGCTTCCACACACATG 540
 QY 645 TTCTGGGCTGCTGGGCTGTGCTTGCACACCTGTCAATGAAAGCGTGGCTTGGAGTGA 704
 DB 541 TTCTGGGCTGCTGGGCTGTGCTTGCACACCTGTCAATGAAAGCGTGGCTTGGAGTGA 600
 QY 705 ATCCGGGACCTTATCTGCTCCAGAGCTGTGCGAAGGATGAACACCTTGGCGTAC 764
 DB 601 ATCCGGGACCTTATCTGCTCCAGAGCTGTGCGAAGGATGAACACCTTGGCGTAC 660
 QY 765 TGTGGACACTGTATGCTTACAGCTGATGATGCTTCCACAGTGTATGACAGAGTGTG 824
 DB 661 TGTGGACACTGTATGCTTACAGCTGATGATGCTTCCACAGTGTATGACAGAGTGTG 720
 QY 825 ACTGTGGCGGTGTACAGCTTCTCTGACTGTAGTTGGGCGGAGTTCTTGAACCCA 884
 DB 721 ACTGTGGCGGTGTACAGCTTCTCTGACTGTAGTTGGGCGGAGTTCTTGAACCCA 780
 QY 885 GCCAAGGCTTACCTGTGCGATGAGCTGGAAGCTGTGCGGCTTCTTCAAGTTCCAG 944
 DB 781 GCCAAGGCTTACCTGTGCGATGAGCTGGAAGCTGTGCGGCTTCTTCAAGTTCCAG 840
 QY 945 TTCTTCTTCTATGTTGGCTGTGCTGAAGTGGCAAGGCTGATCAACCTTGGAGAG 1004
 DB 841 TTCTTCTTCTATGTTGGCTGTGCTGAAGTGGCAAGGCTGATCAACCTTGGAGAG 900
 QY 1005 GATGATGATATTTTGAAGACCACTGATGCTGACAGAAATTTGACAGTGTCCCTGTG 1064
 DB 901 GATGATGATATTTTGAAGACCACTGATGCTGACAGAAATTTGACAGTGTCCCTGTG 960
 QY 1065 GCTGTGATGATGATGACACAGAGCTGTGCTGGAGTGAAGCCGAGATGTACTGGAATAG 1124
 DB 961 GCTGTGATGATGATGACACAGAGCTGTGCTGGAGTGAAGCCGAGATGTACTGGAATAG 1020
 QY 1125 CCGGAGGCAAGGCCCCCTTACAGAGCTGTCCGCGCAAGTTCCGTGAGGCTCTTTATG 1184
 DB 1021 CCGGAGGCAAGGCCCCCTTACAGAGCTGTCCGCGCAAGTTCCGTGAGGCTCTTTATG 1080
 QY 1185 GGCTTCACCTTCAACATCAAGCTGGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAG 1244
 DB 1081 GGCTTCACCTTCAACATCAAGCTGGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAG 1140

QY 1245 GACGAGAGAGATGCTCAAGCTGCGATGTTGGCCGCTTCTTACGAGCTGCAAGTCCCATGAT 1304
 DB 1141 GACGAGAGAGATGCTCAAGCTGCGATGTTGGCCGCTTCTTACGAGCTGCAAGTCCCATGAT 1200
 QY 1305 CACCATCTTCCAGAGGCAAACTCAAGAGCAAACTCACTGTGAGCCCAAGAGGAAATCCCTT 1364
 DB 1201 CACCATCTTCCAGAGGCAAACTCAAGAGCAAACTCACTGTGAGCCCAAGAGGAAATCCCTT 1260
 QY 1365 CTTCAGAGAGGCTTGGCCCAAAACCAAGAGCAGCCAAAGAAAGTTAGGGGCGAGGAA 1424
 DB 1261 CTTCAGAGAGGCTTGGCCCAAAACCAAGAGCAGCCAAAGAAAGTTAGGGGCGAGGAA 1320
 QY 1425 GACAACAGGCTGGAAGCTTAAAGGCTGTGAGAGCCCTTCAAGTCTGGCCCATGTATCAG 1484
 DB 1321 GACAACAGGCTGGAAGCTTAAAGGCTGTGAGAGCCCTTCAAGTCTGGCCCATGTATCAG 1380
 QY 1485 AGGCGAGCTTACTACAGTGTGCGCCACAGAGCAGCCCTCAGCCCACTCCATGTTCTTCCC 1544
 DB 1381 AGGCGAGCTTACTACAGTGTGCGCCACAGAGCAGCCCTCAGCCCACTCCATGTTCTTCCC 1440
 QY 1545 CTGAACCATCAAGGCGCGCTGAAGGCTTCAAGTGTCAAGGCTATGACACCAAGACAA 1604
 DB 1441 CTGAACCATCAAGGCGCGCTGAAGGCTTCAAGTGTCAAGGCTATGACACCAAGACAA 1500
 QY 1605 AGCTTAAAGACTGTGAGTTCTGGGCGCAAGAAAGTTTGAATTGTCTCAGAGAGCGAT 1664
 DB 1501 AGCTTAAAGACTGTGAGTTCTGGGCGCAAGAAAGTTTGAATTGTCTCAGAGAGCGAT 1560
 QY 1665 GGGGCTTGTATGAGACACCCAGAAAGTATCTCAAGTGAAGAGAAACCTGTGAGTTAAC 1724
 DB 1561 GGGGCTTGTATGAGACACCCAGAAAGTATCTCAAGTGAAGAGAAACCTGTGAGTTAAC 1620
 QY 1725 CTGACGATATGCGACAGATCCCGGAAATCACTTCAAGAAACCTTGGAAACATACCA 1784
 DB 1621 CTGACGATATGCGACAGATCCCGGAAATCACTTCAAGAAACCTTGGAAACATACCA 1680
 QY 1785 ACCAATACACACTACACTCAAGATCAATGATCTTATTTGGGCTTGGAAAAACAG 1844
 DB 1681 ACCAATACACACTACACTCAAGATCAATGATCTTATTTGGGCTTGGAAAAACAG 1740
 QY 1845 GATGAAGCACTTCTTAA 1862
 DB 1741 GATGAAGCACTTCTTAA 1758

RESULT 2
 US-09-746-783-3
 Sequence 3, Application US/09746783
 Publication No. US20030044935A1
 GENERAL INFORMATION:
 APPLICANT: Jacobs, Kenneth
 McCoy, John M.
 Lavalley, Edward R.
 Racine, Lisa A.
 Treacy, Maurice
 Spaulding, Vikki
 Agostino, Michael J.
 Howes, Steven H.
 Fechtel, Kim
 TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
 NUMBER OF SEQUENCES: 231
 CORRESPONDENCE ADDRESSES:
 ADDRESSES: Genetics Institute, Inc.
 STREET: 87 Cambridgepark Drive
 CITY: Cambridge
 STATE: MA
 COUNTRY: U.S.A.
 ZIP: 02140
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS


```
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1702)..(1702)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1712)..(1712)
OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-259
```

```
Query Match
Best Local Similarity 26.4%; Score 587.8; DB 17; Length 1717;
Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
```

```
QY 1242 GAGACGAGGAGGATGCTACGCTGATCATTTGGCGCTTCTAGGCTTGCATTCCT 1301
DB 1 GAGACGAGGAGGATGCTACGCTGATCATTTGGCGCTTCTAGGCTTGCATTCCT 60
QY 1302 GATCACCATTCTCCAGGGCAAACTCAAGGACCAACTCTGTGGCCCAAGAGGAAATCC 1361
DB 61 GATCACCATTCTCCAGGGCAAACTCAAGGACCAACTCTGTGGCCCAAGAGGAAATCC 120
QY 1362 CTTCCTCAGAGGGGCTTGCCTCAAAAACCAAGGACGAGCCAAACAGAACTTTAGGGCCAG 1421
DB 121 CTTCCTCAGAGGGGCTTGCCTCAAAAACCAAGGACGAGCCAAACAGAACTTTAGGGCCAG 179
QY 1422 GAAGCAACAAGGCTTGAAGCTTAAGCTGTGAGCGCTTCAAGCTGTGGCCCACTGTAT 1481
DB 180 GAAGCAACAAGGCTTGAAGCTTAAGCTGTGAGCGCTTCAAGCTGTGGCCCACTGTAT 239
QY 1482 CAGAGGCGGCTTACACAGTGCCTCCCAAGAGCGCCCTTCAAGCTGTGGCCCACTGTAT 1541
DB 240 CAGAGGCGGCTTACACAGTGCCTCCCAAGAGCGCCCTTCAAGCTGTGGCCCACTGTAT 299
QY 1542 CCCCTAGAACCATCAGCGCGCTCAAGGCTTCAAGCTGTGAGCGCTTCAAGCTGTGGCCCACTGTAT 1601
DB 300 CCCCTAGAACCATCAGCGCGCTCAAGGCTTCAAGCTGTGAGCGCTTCAAGCTGTGGCCCACTGTAT 359
QY 1602 AAAAGCTTAAGACTGTGATTTCTGGGGCCAAAGAAATTTGATTTGCTCTCAGAGAGC 1661
DB 360 AAAAGCTTAAGACTGTGATTTCTGGGGCCAAAGAAATTTGATTTGCTCTCAGAGAGC 419
QY 1662 GATGGGGCTTGTATGAGACCCAGAAATCTCAAGTGAAGAGAAACCTGTGAGTTT 1721
DB 420 GATGGGGCTTGTATGAGACCCAGAAATCTCAAGTGAAGAGAAACCTGTGAGTTT 479
QY 1722 AACCTGACGATATGCCAGAGATCCCGAAATCACTCAAGAACTTTGGAACATCA 1781
DB 480 AACCTGACGATATGCCAGAGATCCCGAAATCACTCAAGAACTTTGGAACATCA 539
QY 1782 CCAACCAATCACTCACTCACTCAAGATCACTGATCTTATTTGGGCTTTGGAAAC 1841
DB 540 CCAACCAATCACTCACTCACTCAAGATCACTGATCTTATTTGGGCTTTGGAAAC 599
QY 1842 AGGGATG 1848
DB 600 AGGCTG 606
```

RESULT 4
US-10-492-032-3

```
Sequence 3, Application US/10492032
GENERAL INFORMATION:
APPLICANT: Steuernagel, Arnd
APPLICANT: Bronner, Gunter
APPLICANT: Fritsch, Rüdiger
APPLICANT: Eulenber, Karsten
APPLICANT: Closssek, Thomas
TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
Regulation of Energy Homeostasis
FILE REFERENCE: 2923-606
CURRENT APPLICATION NUMBER: US/10/492,032
```

```
CURRENT FILING DATE: 2004-04-08
PRIOR APPLICATION NUMBER: PCT/EP02/11321
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.2
SEQ ID NO 3
LENGTH: 2137
TYPE: DNA
ORGANISM: Homo sapiens
US-10-492-032-3
```

```
Query Match
Best Local Similarity 25.9%; Score 577.2; DB 19; Length 2137;
Matches 793; Conservative 0; Mismatches 338; Indels 3; Gaps 1;
```

```
QY 103 CCATGACCATCACTTACAAAGGCAAGTGTATATCCCGCTTAGGCTCTTCCGCGC 162
DB 210 CCATGACCATCACTTACAAAGGCAAGTGTATATCCCGCTTAGGCTCTTCCGCGC 269
QY 163 TGCTGTGTGTGCGGCGGAGCATCTACAGCTGTATATGAGCGAGTTCTTAATCTTC 222
DB 270 TGCTGTGTGTGCGGCGGAGCATCTACAACTCTGTGTGAGAGCTGTCTTCTTC 329
QY 223 TGCTGTGTGTGTGCGGCGGAGCATCTTATTTATATGCTGTGAGCGAGCAAGCAAGC 282
DB 330 TGCTGTGTGTGTGCGGCGGAGCATCTTATTTATATGCTGTGAGCGAGCAAGCAAGC 389
QY 283 TGATGTTTGAAGAACTGATCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTCCT 342
DB 390 TGATGTTTGAAGAACTGATCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTCCT 449
QY 343 TGCTGTGTGTGTGCGGCGGAGCATCTGTGAGCGCTGTGAGCGAGCAAGCAAGC 402
DB 450 TGCTGTGTGTGTGCGGCGGAGCATCTGTGAGCGCTGTGAGCGAGCAAGCAAGC 509
QY 403 TGCTGTGTGTGTGCGGCGGAGCATCTGTGAGCGCTGTGAGCGAGCAAGCAAGC 462
DB 510 TGCTGTGTGTGTGCGGCGGAGCATCTGTGAGCGCTGTGAGCGAGCAAGCAAGC 569
QY 463 AAGCGGCTGTGTGCGGCGGAGCATCTGTGAGCGCTGTGAGCGAGCAAGCAAGC 522
DB 570 AAGCGGCTGTGTGCGGCGGAGCATCTGTGAGCGCTGTGAGCGAGCAAGCAAGC 629
QY 523 TGCGAGCGGTGTGCGGCGGAGCATCTGTGAGCGCTGTGAGCGAGCAAGCAAGC 582
DB 630 TGCGAGCGGTGTGCGGCGGAGCATCTGTGAGCGCTGTGAGCGAGCAAGCAAGC 689
QY 583 CAGGCTTTATGATCCGCGAGCAAGCAAGCTGTGAGCAAGCAAGCTGTGAGCAAGC 642
DB 690 CAGGCTTTATGATCCGCGAGCAAGCAAGCTGTGAGCAAGCAAGCTGTGAGCAAGC 749
QY 643 TGCTGTGTGTGTGCGGCGGAGCATCTGTGAGCGCTGTGAGCGAGCAAGCAAGC 702
DB 750 TGCTGTGTGTGTGCGGCGGAGCATCTGTGAGCGCTGTGAGCGAGCAAGCAAGC 809
QY 703 GAATCCGGGACCCATCTGTGCTCCAGAGCGCTGTGAGCAAGCAAGCTGTGAGCTC 762
DB 810 GAATCCGGGACCCATCTGTGCTCCAGAGCGCTGTGAGCAAGCAAGCTGTGAGCTC 869
QY 763 AGTGTGACACCTGTATGCTGTGAGCAAGCTGTGAGCAAGCTGTGAGCAAGCTGTGAGC 822
DB 870 AGTGTGACACCTGTATGCTGTGAGCAAGCTGTGAGCAAGCTGTGAGCAAGCTGTGAGC 929
QY 823 TGATGTGTGTGTGCGGCGGAGCATCTTCTGTGAGCTGTGAGCGAGCTGTGAGCAAGC 882
DB 930 TGATGTGTGTGTGCGGCGGAGCATCTTCTGTGAGCTGTGAGCGAGCTGTGAGCAAGC 989
QY 883 CAGCAGAGGCTTACCTGTGAGCAAGCTGTGAGCGCTGTGAGCGAGCTGTGAGCAAGC 942
DB 990 CAGCAGAGGCTTACCAAGCAAGCAAGCTGTGAGCGCTGTGAGCGAGCTGTGAGCAAGC 1049
```

Qy 943 AGTCTTCTTCTATGCTGCTGCTGAGGTGAGAGACAGCTCATCAACCCCTTTGGAG 1002
Db 1050 AGTCTTCTTCTTCTAGCCGGCTGGCTCAGAGTATGCTGAGACAGCTCATCAACCCCTTTGGAG 1109
Qy 1003 AGATGATGATGATTTTGAACCACTGATTTGTCAGACAGAAATTTGCAAGTTCCTCT 1062
Db 1110 AGGACGATGATGATCTTGAACCACTTCTGATGATGATGAACTTCCAGGTGTCACAGC 1169
Qy 1063 TGGCTGATGATGATGATGACACAGAGCTGCTGGATGAGCCGGAGCATTTACTGGAATA 1122
Db 1170 TGGAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1229
Qy 1123 AGCCGAGCCACAGCCCTTACACAGCTGCTTC---CGCCAGTTCCTGAGCTCTCT 1179
Db 1230 CAGCGAGGCTCGGGCCCATACACAGCGGCTACGTCTTCCAGCTGCG3CAGGCTTCTCT 1289
Qy 1180 TTATGGGCTCCACCTTCAACATCAGCTGAAACAAAGAGAGATGAGTTCCAGC 1233
Db 1290 TCCAGGGCTCCACCTTGTGACATCAGCTGCGCCAAAGAGACATGACGTTCCAGC 1343

RESULT 5

US-10-492-032-30
; Sequence 30, Application US/10492032
; Publication No. US20050049212A1
; GENERAL INFORMATION:
; APPLICANT: Steuernagel, Arnd
; APPLICANT: Bromer, Gunter
; APPLICANT: Fritsch, Rüdiger
; APPLICANT: Eulenberg, Karsten
; APPLICANT: Closssek, Thomas
; TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
; TITLE OF INVENTION: Regulation of Energy Homeostasis
; FILE REFERENCE: 2923-606
; CURRENT APPLICATION NUMBER: US/10/492, 032
; PRIOR FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: PCT/EP02/11321
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: EP01124059.5
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 1530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-492-032-30

Query Match 25.9%; Score 576.8; DB 19; Length 1530;
Best Local Similarity 70.0%; Pred. No. 2,1e-154;
Matches 792; Conservative 0; Mismatches 337; Indels 3; Gaps 1;

Qy 105 ATGACCATGACTTACACAGCAAGCTGCTATATGCTGCTTCTCCGCTG 164
Db 1 ATGACCGCTGCTACACAGCCCGAGTGGCAAGCCGCTTGGTGGCTTCTCCAGCTG 60
Qy 165 CTGCTGCTGCTGGCGGGGAGAGATCTCAAGCTGCTATATGCGAGTTCTTATCTCTG 224
Db 61 CTGCTACTGTGGCGGTGGAGATCTCAAACTCTGTGGCAGAGCTGCTTGTCTCTT 120
Qy 225 CTGCTACTGATCATCCGCTTATTTATATGAGCTGCGCTTCAAGAAACAGAGCTG 284
Db 121 GGGTTCATAGCGGCTGAGTGCCTTACCGCTTTGTGCTGACCGAAGGGCAGAGCGC 180
Qy 285 ATGTTTGAAGATGACTCTGTATTTGACAGAGCTACATCCAGCTATCCCAATTTCTTC 344
Db 181 TACTTCGAGAGCTGTGATTTATTTATGACAGATATGCAACCTCATCTCTGCTCTTC 240
Qy 345 GTGCTGGCTTCTTACAGTACGCTGCTGACCCGCTGCTGAGAAACAGTACAGAACTG 404
Db 241 GTGCTGGCTTCTTATGATACGCTGCTGAGAAACGCTGCTGAGCACTTACCTATGACTG 300
Qy 405 CCGTGGCCGACCGCTCATGAGCTGTGCTGGGCTTCTGCAAGCAAGAGAGAGCA 464

Db 301 CCGTGGCCGACCGCTCATGCTGCTGAGTGGCGGACCGGACAGAGAGAGAGAGAGAGAG 360
Qy 465 GCGCGGCTGCTGGCGGCGGACGCTCATCCGCTACGCGCAACCTGGGCAACGCTCATCTG 524
Db 361 GCGCGGCTTACCGGCGGACCTCATGCTGCTACCGAGGCTCTGGCGGTCTTACTCTG 420
Qy 525 CCGAGCTACAGACCGGAGTCTACAGAGCTTCCCGAGCGGCGGACGCTGCTGCAAGCA 584
Db 421 CCGTCCGTCAGACCGGCGGCTTCAAGCGCTTCCCGACATAGACAGTGTGAGAGCT 480
Qy 585 GCGTTTATGACTCCGAGCAACACAGAGCTTGAAGAACTGACCTTACACACACATG 644
Db 481 GGGTTATGACCCCGAGAGAGCGCAAGAACTTGAAGAACTGAACTCATCTTACAAACAG 540
Qy 645 TTCTGGGTCCTGCTGCTGCTGCTTCCAACTGTCATATGAAAGGCTGCTTGGAGCTCA 704
Db 541 TACTGGGTCCTGCTGCTGCTGCTTCCAACTGCGGCGACAGGCGCAAGCGGAGGCGC 600
Qy 705 ATCCGAGACCTTATCTGCTTCAAGAGCTGCTGCAAGAGATGAACACTTGTACTTCA 764
Db 601 ATCCGAGACCAAGAGCGGCTTATGAGCTGCTGCTGCAAGAGATGAAGTTTTCGGGCG 660
Qy 765 TGTGACACCTGATAGCTTACAGCTGATTAATGATCCACTGCTGATATACAGAGTGTG 824
Db 661 TGTGAAATGCTTCTTCACTATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 825 ACTGTGCGGTGATACAGCTTCTTCTGATCTGCTTATGATGATGATGATGATGATGATG 884
Db 721 ACCATGCACTGATACAGCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATG 780
Qy 885 GCCAAGGCTTACCTGCGCATAGAGCTGATGATGATGATGATGATGATGATGATGATGATG 944
Db 781 GCTAGAGGTTTCAAAAGACACAGACTTATGATGATGATGATGATGATGATGATGATGATG 840
Qy 945 TTCTTCTTCTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1004
Db 841 TTCTTCTTCTTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy 1005 GATATGATGATTTTGAAGCAACATGATGATGATGATGATGATGATGATGATGATGATG 1064
Db 901 GACATGATGATCTTGTGAGCACTTCTGATGATGATGATGATGATGATGATGATGATGATG 960
Qy 1065 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1124
Db 961 GCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
Qy 1125 CCGAGCCAGAGCCCGCTTACACAGCTGCTTC---CGCCAGTTCGCTGAGCCTCTT 1181
Db 1021 GCGAGGCTGCGCGCCCATACACAGCGGCTATCTGCTTCCAGCTGCGGACGCTTCTTC 1080
Qy 1182 ATGGGCTTCCACTTCAACATCAGCTTGAACAAAGAGATGAGTTCCAGC 1233
Db 1081 CAGGGCTTCACTTGTGACATCACGCTGCGCAAGAGACATGACGTTCCAGC 1132

RESULT 6

US-10-108-260A-1299
; Sequence 1299, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1 full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108, 260A
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1299
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1299

Query Match 20.6%; Score 459.2; DB 17; Length 2500;
 Best Local Similarity 63.5%; Pred. No. 17e-120;
 Matches 701; Conservative 0; Mismatches 403; Indels 0; Gaps 0;

QY 105 ATGACCATCTTACACAGCCCAAGTGGTAAATGCCGCTAGAGCTCTCTCCGCTG 164
 DB 229 ATGACTGTCTACTTCTCCAGTAAAGTAGCAATCTTTTGGATTCTAGTTA 288
 QY 165 CTGCTGTGGGGGAGCATCTCAAGCTCTTATGCGAGTTCTTATCTTCTG 224
 DB 289 CTCTCAAGTGGAGGAGCATCTCAAACTACTAGCAAGGAATTTTGTGTTGCT 348
 QY 225 CTCTGCTACTACATCATCCGCTTATTTATGAGCTGGCCCTCAGGAAAGAACAGCTG 284
 DB 349 GTTCTTATACAGCAATTAAGTTTGTGTACAGATTCTTCTTACAGAGATCCAAAAGCT 408
 QY 285 ATGTTGAGAAATGACTCTGTATTCGACAGCTTACATCCAGTCAATCCCATTTCTTC 344
 DB 409 TACTTGAATAATTAATTAATTAATCTGTACAGATATGCTGAACAAATTCAGTAACTTT 468
 QY 345 GTGCTGGGCTTCTACGTGACGTGTGTCGACCCGCTGGTGAACCAAGTAAAGCTG 404
 DB 469 GTGCTTGGGTTTATATTTACTGTGATGAAACCATGTGGAGACAGTTGTGATTTG 528
 QY 405 CCGTGCCCGAGCCGCTCATGAGCTGTGTGGGCTTCTGCGAAGGAGAGAGAGCA 464
 DB 529 CCGTGCCAGACAGGCTAATGTTCCTCATCTAGAGAGTTTCAGGAAGAGAGAGAGC 588
 QY 465 GGCAGGCTCTGGGGGAGCGCTCATCCGCTTACGCAACTGGGCAAGTGTCTCTG 524
 DB 589 GGGCGGCTCTTGAAGAGCGCTGATGCGCTAGTCAATCTCATCTCTGCTCAATCTT 648
 QY 525 CCGAGCGTCAGACCGGAGCTTACAAAGCGCTTCCCGAGGCGCCAGACACTGTGTGAAGCA 584
 DB 649 CCGTGGTGAGCACTGTGTGTACAAAGATTTCACAAATGAGCACTAGGTTGAAGCA 708
 QY 585 GGCCTTATGACCTCGGAGCAACAGAGAGTTGGAAAGCTTACAGCTTACAGCAACTG 644
 DB 709 GCTTTATGACAAAGATGAAAGAAATTTTCAACCACTCAAGTCTCTCACTGAGAA 768
 QY 645 TTCTGGGTCCTGGGCTGTGTGTCCAACCTGTCAATGAGAGCGTGGCTTGAAGTGA 704
 DB 769 TATGAGGTTCCATTCATCTGGTTTGAATCTGTGACATTAAGCCGGAATGAAGTGA 828
 QY 705 ATCCGAGACCTTATCTGCTCTCAAGAGCTGTGAACGAGATGAACACTTGGCTGAG 764
 DB 829 ATCAAGACAGTGTGATCTGCAATCATTTGATGACTGAATGAATGATACCGCTTGG 888
 QY 765 TGTGACACCTGTATGCTTACAGATGATGATATCCCACTGTGTATACAGAGTGTG 824
 DB 889 TGCAGCCTTATTCGTTATGACTGGTGTGAATCCGCTGTTTACACCCAGGTTGTC 948
 QY 825 ACTGTGGGCTGTACAGCTTCTTCTGACTGTCTAGTTGGGCGGAGTTTGAAGCCA 884
 DB 949 ACTGTGTGTCTATCTTCTTCTTGTGCGCTGATGTGAAGCGCAGTTTGTGATCC 1008
 QY 885 GCAAGAGCTTACCTGTGCAATGAGTGAAGTGAAGAGCTCATCAACCCCTTGAAGAG 944
 DB 1009 ACCAAAGCTTACGAGGAGCACTGTGATCTTTCATCTTCACTTCAACCTCTTCA 1068
 QY 945 TTCTTCTTATGTTGTGCTGTGAGAGTGAAGTGAAGAGCTCATCAACCCCTTGAAGAG 1004
 DB 1069 TTCTTCTTATGAGAGATGGCTTAAAGTGAAGAGAGCTTATCAACCTTGTGAAGAA 1128
 QY 1005 GATGATATGATTTTGAAGCAACATGAGATGTGACAGAAATTTGAGAGTGTCTGTTG 1064
 DB 1129 GATGATATGATTTTGAAGCACTGTGATGTGACAGAAATTTGAGAGTGTCTGTTTAA 1188
 QY 1065 GCTGTGATGAGATGACAGAGACCTGCTCGATGAGAGAGGAGCATGTACTGGAATAG 1124
 DB 1189 GCTGTGAGCAATGACATGAGCTTACCAAGATGAGAGAGCAATTTACTGAGAGCAT 1248

QY 1125 CCCAGCCACAGCCCCCTTACAGAGCTGTTCGCCCCAGTTCCGTCGAGCTCTTATG 1184
 DB 1249 TCTGTGCTGCGCCACCAATACATGAGCTGTCTGACTTACATGACCTTATTTCTG 1308
 QY 1185 GAGTCACCTTCAACATCAGCTG 1208
 DB 1309 GGGTACAGTTCAGATGGGCTG 1332

RESULT 7

US-10-492-032-7
 ; Sequence 7, Application US/10492032
 ; Publication No. US20050049212A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Steuernagel, Arnd
 ; APPLICANT: Bronner, Gunter
 ; APPLICANT: Fritsch, Rudiger
 ; APPLICANT: Eulenberg, Karsten
 ; APPLICANT: Ciossek, Thomas
 ; TITLE OF INVENTION: Receptorin and Receptorin Homologous Proteins Involved in the
 ; FILE REFERENCE: 2923-606
 ; CURRENT APPLICATION NUMBER: US/10/492,032
 ; PRIOR FILING DATE: 2004-04-08
 ; PRIOR APPLICATION NUMBER: PCT/EP02/11321
 ; PRIOR FILING DATE: 2002-10-09
 ; PRIOR APPLICATION NUMBER: EP01124059.5
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 7
 ; LENGTH: 1422
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-492-032-7

Query Match 20.0%; Score 446.2; DB 19; Length 1422;
 Best Local Similarity 63.8%; Pred. No. 6.7e-117;
 Matches 751; Conservative 0; Mismatches 378; Indels 48; Gaps 3;

QY 105 ATGACCATCTTACACAGCCCAAGTGGTAAATGCCGCTAGAGCTCTCTCCGCTG 164
 DB 1 ATGACGTTTCAATACCTCTCAAGTGGGAGGCCCTTGGAGTTCTCTGCGCTG 60
 QY 165 CTGCTGTGCTGGGGGAGCATCTTACAAAGCTGTATATGAGAGTCTTAACTTCTCTG 224
 DB 61 CTCTCCGCTGAGGGGAGGAGCATCTACAGCTCTCTTCAAGAAATTTCTCTTGTGG 120
 QY 225 CTCTGCTACTATCATATCCGCTTATTTATAGCTGGCCCTCAGCGAAGAAACAAGCTG 284
 DB 121 GCTTGTATGCTGTGCTTATGACTACCTACCGGCTGTGTGACCCAGAGCAAGATAC 180
 QY 285 ATGTTGAGAAATCTGATTTGATTTGCAAGCTTATCCAGCTTATCCCATTTCTTC 344
 DB 181 GTGTATGCTCAGTGGCCGAGTCAACCGCTCAGAGACTTCAATTTCTTGTCTTT 240
 QY 345 GTGCTGGCTTCTACGTGACGCTGTGTCGTCACCGCTGTGGAACCAAGTACAGAACTG 404
 DB 241 GATTTGGGTTCTTATGATCTGTGTGTGAACCGCTGTGTGTGTCCAGTACAAAGATC 300
 QY 405 CCGTGGCCCGAGCCGCTTATAGAGCTGTGTGGGCTTGTGGAAGCAAGAGAGCA 464
 DB 301 CCGTGGCCAGACAGCTGATGTGTATCTGTGCTAGCGGTGACAGCGGTGACAGCGG 360
 QY 465 GAGCGGCTGTGCGGGCGAGCTCATCGCTAGGCGCAACTGGGCAACAGTGTCTATCTG 524
 DB 361 GGGCGGCTGTGCGCGGAGCTTATCGCTTACGCAACTGGGCTCGGTGTGTGTGTG 420
 QY 525 CGAGGCTGACAGCGAGTCTTACAGCGCTTCCCAAGCGCCAGCACTGTGTGCAAGCA 584
 DB 421 CGTCTGTGAGCAACCGCGTGTCTTAAAGCGCTTCCCAAGTGTGAGAGAGAGTGTGAGCA 480
 QY 585 GGCCTTATGACTCGGGAGAAACAGAGCAATGAGAACTGAGCTTACCAACAAATG 644

Db 481 GGTTCATGTCCAGAGAGAGAGAAAAGTTTGAGAGCTGAAATCCGACTTCAACAG 540
Qy 645 TTTCGGGCGCCCTGGGTGTGTTGGCCAACTGTCAATGAAGGCGTGTGAGGTGGA 704
Db 541 TACTGGGTCCTCGTGTGTGTTCAACCACTGGGGCCCGGAGGAGGAGCGGGGA 600
Qy 705 ATCCGGGACCTTATCTGCTCCAGAGCCTGTGTAAGAGATGAACAACCTTGCTACTGAG 764
Db 601 ATAGTGAACGATATGCTCTGTCTACTTTTGGAGAGCTGAACAGTATACAGGCCAAG 660
Qy 765 TGTGGAACCTGTATGCTTACAGATGATATGATATCCCACTGTGTATATACAAGGTGTG 824
Db 661 TGACAGATGATATTCATATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 825 ACTGTGGGTGTATAGCTTCTTCTGACTGTGTATGTGAGGCGGCAAGTTTGAACCA 884
Db 721 ACCATAGCCGTCTACTTCTTCTTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy 885 G-----CCAGGCGCTTACCCTGGCC 903
Db 781 GAGGAGGAGGCTGCCAACTCAAGAGCTTGTGAAGCCAGGCGAGAGCCAGCCCGAGCC 840
Qy 904 AT---GAGCTGAGCTGTGTGTGCTGTCTTCAAGTCTGTCAAGTTCTTCTATGTT 959
Db 841 CTGGAGAGCCGGAGCATGTAGTGTCTCTCACTGTGTGAGTTCTTCTATGCT 900
Qy 960 GCGTGGCTGAAGGTGTGAGAGAGAGCTATCAACCCCTTTGAGAGAGATGATGATTTT 1019
Db 901 GCGTGGCTCAAGGTGTGTGAGAGAGATCAATCAACCATTTGTGTGAGAGATGAGAGACTTT 960
Qy 1020 GAGACCAATGATTTGTGAGAGAGATTTGAGAGGTGCTGTGTGCTGTGTGATGATGATG 1079
Db 961 GAGACCAATGATTTGTGAGAGAGATTTGAGAGGTGCTGTGTGCTGTGTGATGATGATG 1020
Qy 1080 CACCAGGACCTGCTCGGATGAGAGCGGACATGTATGATGATGATGATGATGATGATG 1139
Db 1021 TACCAGAACTTCCCGCGCTGAGAGAGCAAGTATGATGATGATGATGATGATGATG 1080
Qy 1140 CCCTTACAC---AGCTGTCTCCGCCCAAGTTCCGTGAGGCTCTTATATAGCTTCAACCTTC 1136
Db 1081 CCCTTACACTGTGTGAGAGAGCGGCGGAGTCTGTGCGGCTCTTATATAGCTTCAACCTTC 1140
Qy 1197 AACATCAGCCTGAACAAAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1233
Db 1141 AACCTGGCATGAG 1177

RESULT 8
US-10-492-032-5
Sequence 5, Application US/10492032
Publication No. US20050049212A1
GENERAL INFORMATION:
APPLICANT: Steuermagel, Arnd
APPLICANT: Brionner, Gunter
APPLICANT: Fritsch, Rudiger
APPLICANT: Eulenberger, Karsten
APPLICANT: Closser, Thomas
TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
FILE REFERENCE: 2923-606
CURRENT APPLICATION NUMBER: US/10/492,032
CURRENT FILING DATE: 2004-04-08
PRIOR APPLICATION NUMBER: PCT/EP02/11321
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.2
SEQ ID NO 5
LENGTH: 2028
TYPE: DNA
ORGANISM: Homo sapiens

US-10-492-032-5
Query Match 19.0%; Score 423; DB 19; Length 2028;
Best Local Similarity 62.2%; Pred. No. 3.7e-110;
Matches 697; Conservative 0; Mismatches 400; Indels 24; Gaps 1;
Qy 105 ATGACATCACTTACAGAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATG 164
Db 1 ATGACATCACTTACAGAGCAAGGATGATGATGATGATGATGATGATGATGATGATGATG 60
Qy 165 CTGCTGTGCGCGGCGAGATCTTACAGAGTGTATATGAGCGAGTTCTTATCTTCTG 224
Db 61 CTCTCAAGTGAAGAGCAGATCTTACAACTGTATGAGGAGATTTATGATTTTGTCT 120
Qy 225 CTCTCACTACATCATCCGCTTATTTATGAGTGTGCGCTTCAAGAGAGCAAGCTG 284
Db 121 GTCTTTATACAGAAATGATTTGTATACAGATTTATCTTACAGAGTGTCAAAACGT 180
Qy 285 ATGTTGAGAACTGACTGTGTATGAGAGAGCTATCCAGCTATCCCATTTCTTC 344
Db 181 TACTTGAATAATTTATCAATTTACTGTAGACAGATATGCTGAACAAATTCAGTAACTTT 240
Qy 345 GTGCTGAGCTTCTATGAGAGCTGTGTGATACCGCTGTGGAACCAATGAGAGACTG 404
Db 241 GTGCTGAGCTTCTATGATCTGTGTGATGATGATGATGATGATGATGATGATGATG 300
Qy 405 CCGTGGCCGACCGCTCATGAGCTGTGTGTGCGGCTTGTGCGAAGCAAGAGAGCA 464
Db 301 CCGTGGCCGACCGCTCATGAGCTGTGTGTGCGGCTTGTGCGAAGCAAGAGAGCA 360
Qy 465 GAGCGGCTGTGCGCGCGCAGCTCATCCGATCCGATCCGATCCGATCCGATCCGATCCG 524
Db 361 GAGCGGCTGTGCGCGCGCAGCTCATCCGATCCGATCCGATCCGATCCGATCCGATCCG 420
Qy 525 CGAGAGCTCAGACCGAGCTGTACAGAGCTTCCCGAGCGCCAGCACTGTGTGAGCA 584
Db 421 CGCTGTGTGAGACTGTGTGTACAAAGATTTCCCAATGAGACAGTGTGTGAGCA 480
Qy 585 GCGTTTATGACTCGGAGAGAACCAAGAGTGTGAGAACTGAGAGCTTACCAACATG 644
Db 481 GGTTTATGACAAAGATGAAAGAGAAATTTATCAACCACTCAAGTCTTCTCATCTGAA 540
Qy 645 TTTCGTGCGCCGTGGGTGTGTTGCAACCTGTCAATGAAAGCGTGTGAGTGTGGA 704
Db 541 TATTGGTTCATTTCACTGTGTGTGAAATCTTGCAATGAAAGCGGAAATGAAAGTGA 600
Qy 705 ATCCGGAGCCTTATCTGCTTCCAGAGCTGTGAAAGAGATGAACACTTGTGATCTAG 764
Db 601 ATCAGAGACAGTGTGTGATCTGCAATCATGTATGATGATGATGATGATGATGATG 660
Qy 765 TGTGGAACCTGTATGCTTACAGATGATGATGATGATGATGATGATGATGATGATGATG 824
Db 661 TGAGGCTCTTATTCGGTTTATGACTGGTTGAGATTCGCTGTGTATACACCAAGTGTG 720
Qy 825 ACTGTGCGGTGTACAGCTTCTTCTGACTGTGTATGTGAGGCGGAGTTTGAACCA 884
Db 721 ACTTGTGCTGTATGATCTTCTTCTTGTGCTGTGTATGAGCGCAGTTTGTGATGCC 780
Qy 885 GCCAAGGCTTACCTGTGCAATGAGCTGAGACTGTGTGCGCGCTTCAAGTGTCTGAG 944
Db 781 ACCAAGGCTTACCGAGGAGATGATGATGATGATGATGATGATGATGATGATGATG 840
Qy 945 TTCTTCTTATGTTGGCTGTGAG-----GTGCGAGAG 980
Db 841 TTCTTCTTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Qy 981 CAGCTCATCAACCCCTTGTGAGAGATGATGATGATGATGATGATGATGATGATGATG 1040
Db 901 CAGCTTATCAACCTTTTGTGAGAGATGATGATGATGATGATGATGATGATGATGATG 960
Qy 1041 AGGAATTTGAGAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1100
Db 961 AGAATTTGAGAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020

QY 1101 GAGCCGACATGATCTGGAATAGCCCGAGCCAGCCCCCTCAACAGCTGCTCCGCC 1160
DB 1021 AAGAGGACATTTACTGGACGATTTCTGCTGCTGCGCCACCAATGACATTTGGAGCTGCT 1080
QY 1161 CATTCCGCTGAGCCCTCTTTATGGGCTCCACCTTCAACAT 1201
DB 1081 GACTACTGCATACCTTCATTTCTGGGGTCAACAGTCCAGAT 1121

RESULT 9

US-10-071-766-10/c
; Sequence 10, Application US/10071766
; Publication No. US20020192678A1
; GENERAL INFORMATION:
; APPLICANT: Huel-Mei Chen
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
; FILE REFERENCE: PA-0043 US
; CURRENT APPLICATION NUMBER: US/10/071,766
; CURRENT FILING DATE: 2002-02-07
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020192678A1 1138151.2
US-10-071-766-10

Query Match 15.7%; Score 350.4; DB 13; Length 1326;
Best Local Similarity 99.7%; Pred. No. 2.1e-89;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1850 AGCATTCTTAACTCTTCTTAAATGGGAGTCTTGGCCAGCCAGTCTCAGCTGTGT 1909
DB 1326 AGCATTCTTAACTCTTCTTAAATGGGAGTCTTGGCCAGCCAGTCTCAGCTGTGT 1267
QY 1910 GTACACGAGCAGCAGTATGTCAGTCAAGCAGTACAGTCTTCCACCTGATGAT 1969
DB 1266 GTACACGAGCAGCAGTATGTCAGTCAAGCAGTACAGTCTTCCACCTGATGAT 1207
QY 1970 TCTTAAACAGCCTGATCAATGATGTTAGTTAAATGAAATCCAGACTGTTAG 2029
DB 1206 TCTTAAACAGCCTGATCAATGATGTTAGTTAAATGAAATCCAGACTGTTAG 1147
QY 2030 CCTTAAATGCTTTATTTATTTATTTAAATGTTAAATGTTAAATGTTAAATGTT 2089
DB 1146 CCTTAAATGCTTTATTTATTTATTTAAATGTTAAATGTTAAATGTTAAATGTT 1087
QY 2090 AACTGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2149
DB 1086 AACTGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1027
QY 2150 CACCTTAAATGCTTTATTTATTTATTTAAATGTTAAATGTTAAATGTTAAATGTT 2201
DB 1026 CACCTTAAATGCTTTATTTATTTATTTAAATGTTAAATGTTAAATGTTAAATGTT 975

RESULT 10

US-09-768-826-16
; Sequence 16, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PF512P1
; CURRENT APPLICATION NUMBER: US/09/768,826
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148,759

; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1135)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1148)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1174)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1181)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1209)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1229)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1266)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1290)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1285)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1287)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1295)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1305)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1324)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1339)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1341)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1343)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1345)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1348)
; OTHER INFORMATION: n equals a,t,g, or c
; OTHER INFORMATION: n equals a,t,g, or c
US-09-768-826-16

Query Match

Best Local Similarity 9.7%; Score 215.4; DB 9; Length 1350;
Matches 420; Conservative 0; Mismatches 211; Indels 48; Gaps 3;

QY 603 GAACACAGCACTTGGAGAACTGAGCTTACCAACAACATGTTCTGGGTGCCCTGGGTG 662

Query Match 9.7%; Score 215.4; DB 18; Length 1350;
Best Local Similarity 61.9%; Pred. No. 1.3e-50;
Matches 420; Conservative 0; Mismatches 211; Indels 48; Gaps 3;

QY 603 GAACACAGCAGTGTGAGAAAGTGAAGCTTACACACAGATGTTCTGGGCTGGGCTG 662
DB 23 GAGCGCTGGGGGTTTGAAGAGCTTAATCCGACTTCAACAGTACTGGGCTGGCTG 82
QY 663 TGGTTTCCCACTGTATGAGAGGCGTGGCTTGAAGTGAATCCGGGACCTTATCTG 722
DB 83 TGGTTCACCAACCTGGGGGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 142
QY 723 CTCCAGAGCTGTGAGACAGATGAGACCTTGGGATCTAGGTGAGACCTGTATGCT 782
DB 143 CTCTGTACTTCTTGGAGAGCTGAGACAGTACCGAGCCAGGAGGAGGAGGAGGAGG 202
QY 783 TAGCAGCTGATTTAGTATCCCACTGGTGTATACACAGTGGTGAAGTGGGCTGTA 842
DB 203 TATGACTGATGACATCCCTCTGCTTACACCAAGGAGGAGGAGGAGGAGGAGGAG 262
QY 843 TTCTTCTGACTGTGATGTTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 885
DB 263 TTCTTGGCCCTCTCTGTTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 917
QY 886 -----CCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 917
DB 323 CCTCAGAGAGCTTCTGAGAGCAGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAG 382
QY 918 GTTGTGCGGCTCTTCAAGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 977
DB 383 TAGGCTCTCTCAACCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 442
QY 978 GAGCAGCTCATCAACCTTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1097
DB 443 GAACAGATCATCAACCTTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 502
QY 1038 GACAGAGATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1097
DB 503 GACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 562
QY 1098 ATGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1154
DB 563 GCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 622
QY 1155 TCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1214
DB 623 GCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 682
QY 1215 GAGGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1233
DB 683 GACCGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 701

RESULT 12
US-09-814-353-19504/c
; Sequence 19504, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-0068
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR APPLICATION NUMBER: US 2001-03-21
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940

;; PRIOR FILING DATE: 2000-06-15
;; PRIOR APPLICATION NUMBER: US 60/216,820
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: US 60/220,661
;; PRIOR FILING DATE: 2000-07-25
;; PRIOR APPLICATION NUMBER: US 60/257,672
;; PRIOR FILING DATE: 2000-12-21
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 19504
;; LENGTH: 620
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: 1
;; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19504

Query Match 9.6%; Score 213; DB 10; Length 620;
Best Local Similarity 66.4%; Pred. No. 4.1e-50;
Matches 306; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 103 CCATGACCATCATTTACACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 162
DB 464 CCAATGAGGTTTATACATCTTCAAGTGGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 405
QY 163 TGTGCTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 222
DB 404 TGTCTTCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 345
QY 223 TGTCTGCTATCATCATTCGCTTATTTATTTATTTATTTATTTATTTATTTATTTAT 282
DB 344 GGGCTTGTACGCTGTGCTTACATCACCTACCGGCTGCTGTGAGGAGGAGGAGGAG 285
QY 283 TGAATTTGAGAACTGATCTGTATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 342
DB 284 AGGTGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 225
QY 343 TGTGCTGTGCTTCTTACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 402
DB 224 TTGTATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 165
QY 403 TCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 462
DB 164 TCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 105
QY 463 AAGCGGCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 522
DB 104 GGGGCGGCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 45
QY 523 TGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 563
DB 44 TGGCTCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4

RESULT 13
US-10-198-846-11070
; Sequence 11070, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steilmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR APPLICATION NUMBER: 2002-07-18
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/306,220

NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1287, 1288, 1289, 1290, 1291, 1292
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

Query Match 8.5%; Score 189.2; DB 14; Length 1292;
Best Local Similarity 62.4%; Pred. No. 4.2e-43;
Matches 296; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

Qy 349 TGGGCTTACGTGACGCTGTGTGACCCCGTGTGAACCAAGTACAGAACTGGCGT 408
Db 318 TAGGGTTTATGTACTCTGTGTGTGAACCGATGGTGAACCAAGTTTGTGAATTTGCCCT 377
Qy 409 GGGCCGACCGGCTTACGTGACGCTGTGTGACCCCGTGTGAACCAAGTACAGAAAGGCC 468
Db 378 GGGCAGACAGGCTATGTCTTCTCATCTTAGAGTTCACGGAAGCAAGACGACGGGC 437
Qy 469 GGTGCTGCGGCGACGCTCATCCGCTACGCAACCTGGCAACGTCCTCATCTGCGCA 528
Db 438 GCTGCTTAAAGACGCTGATGCGCTACGTCATCACTTCCCTGCTCATCTTGGCT 497
Qy 529 GCGTCAGACCGCAGCTTACAAAGCGCTTCCCGAGCGCCAGCACTGTGTGAAGAGGCT 588
Db 498 CGGTAGAGACTGTGTGTACAAAGATTCCCAATGAGACCAAGTGTGAAGAGGTT 557
Qy 589 TTAAGACTCCCGGACAAACAAAGAGTGTGAAGAACTAGCTTACCAACAAATGTTCT 648
Db 558 TTAAGACAAAGATGAAGAAATTAATTAACCAACTTAAGTCTCTCATCTGAATATTT 617
Qy 649 GGTGCGCTGAGTGTGTGTGCAACCTGTCAATGAAGCGGTGTGAAGTCAATCC 708
Db 618 GGGTTCCTTATCTGCTGTGTGAATCTTGCAACTGAAGCCCGAATGAAGTGAATCA 677
Qy 709 GGAACCTTATCTGTGTGACAGGCTGTGAAGAGATGAACCTTGTGTACTCAGTGTG 768
Db 678 GAGACAGTGTGATCTGAATCATGTATGATGATGAATGAATGATGATGATGATGATG 737
Qy 769 GACACCTTATGCTTACAGACCTGTATGATGATGATGATGATGATGATGATGATG 822
Db 738 GCTCTTATGCTGTTATGATGATGATGATGATGATGATGATGATGATGATGATG 791

RESULT 14
US-09-880-107-2174/c
Sequence 2174, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Iwe
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
PRIOR FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2174
LENGTH: 1198
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020142981A1 L20941
US-09-880-107-2174

Query Match 7.9%; Score 175; DB 9; Length 1198;
Best Local Similarity 100.0%; Pred. No. 4.8e-39;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2027 CAGCCTTAAATGCTTTTATTCATATAAACTGTGAAGCTAGACTGAACCAATTGGAAACA 2086
Db 1198 CAGCCTTAAATGCTTTTATTCATATAAACTGTGAAGCTAGACTGAACCAATTGGAAACA 1139
Qy 2087 TTTAATCAGACTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGATCCAGACAG 2146
Db 1138 TTTAATCAGACTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGATCCAGACAG 1079
Qy 2147 CCACACCTTAAATACGCTTAACTAATGATTAATTAATTAATTAATTAATTAATTAAT 2201
Db 1078 CCACACCTTAAATACGCTTAACTAATGATTAATTAATTAATTAATTAATTAATTAAT 1024

RESULT 15
US-10-027-632-134530/c
Sequence 134530, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 134530
LENGTH: 578
TYPE: DNA
ORGANISM: Human
US-10-027-632-134530

Query Match 7.2%; Score 161.4; DB 13; Length 578;
Best Local Similarity 99.4%; Pred. No. 2.6e-35;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 582 GCAGGCTTATGACTCCGGGACAAACAGAGTGTGAAGAACTGAGCTTACCAACAC 641
Db 183 GCAGGCTTATGACTCCGGGACAAACAGAGTGTGAAGAACTGAGCTTACCAACAC 124
Qy 642 ATGTTCTGGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701
Db 123 ATGTTCTGGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 64
Qy 702 CGAATCGGAGCCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
Db 63 CGAATCGGAGCCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 21

Search completed: March 26, 2005, 22:07:33
Job time: 1249.65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 22:47:44 ; Search time 355.338 Seconds
(without alignments)
10264.201 Million cell updates/sec

Title: US-09-622-964A-2

Perfect score: 2229
Sequence: 1 cagggaggtccaccagccta.....aaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCUTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2167.8	97.3	2171	US-09-949-016-4901	Sequence 4901, Ap
2	642.2	28.8	7609	US-09-949-016-16644	Sequence 16644, A
3	642.2	28.8	18075	US-09-949-016-16643	Sequence 16643, A
4	358.4	16.1	1462	US-09-949-016-4902	Sequence 4902, Ap
5	188.4	8.5	601	US-09-949-016-173056	Sequence 173056,
6	161.4	7.2	601	US-09-949-016-173051	Sequence 173051,
7	161.4	7.2	601	US-09-949-016-173052	Sequence 173052,
8	128	5.7	601	US-09-949-016-173045	Sequence 173045,
9	102.4	4.6	419	US-09-270-767-735	Sequence 735, App
10	102.4	4.6	419	US-09-270-767-16017	Sequence 16017, A
11	71.4	3.2	601	US-09-949-016-173053	Sequence 173053,
12	70.4	3.2	601	US-09-949-016-173054	Sequence 173054,
13	66	3.0	601	US-09-949-016-173053	Sequence 173053,
14	66	3.0	601	US-09-949-016-173054	Sequence 173054,
15	66	3.0	18075	US-09-949-016-16643	Sequence 16643, A
16	49.2	2.2	601	US-09-949-016-173056	Sequence 173056,
17	49.2	2.2	1462	US-09-949-016-4902	Sequence 4902, Ap
18	49.2	2.2	2171	US-09-949-016-4901	Sequence 4901, Ap
19	49.2	2.2	7609	US-09-949-016-16644	Sequence 16644, A
20	47.4	2.1	1937	US-09-647-143-1	Sequence 1, Appl1
21	46.4	2.1	1101	US-09-902-540-8038	Sequence 8038, Ap
22	46.4	2.1	6808	US-09-902-540-819	Sequence 819, App
23	45.8	2.1	1048	US-09-488-847-38	Sequence 38, Appl
24	45.4	2.0	615	US-09-252-991A-10075	Sequence 10075, Ap
25	45.4	2.0	1116	US-09-252-991A-9937	Sequence 9937, Ap
26	45.4	2.0	1926	US-09-252-991A-9758	Sequence 9758, Ap
27	44.8	2.0	2341	US-09-520-781-9	Sequence 9, Appl1

28	44.8	2.0	2607	US-09-520-781-11	Sequence 11, Appl
29	43.2	1.9	1356	US-09-520-781-31	Sequence 31, Appl
30	43.2	1.9	2737	US-09-902-540-3798	Sequence 3798, Ap
31	43.2	1.9	2823	US-09-252-991A-7100	Sequence 7100, Ap
32	43.2	1.9	3456	US-09-252-991A-7025	Sequence 7025, Ap
33	43.2	1.9	18551	US-09-902-540-1187	Sequence 1187, Ap
34	42.6	1.9	558	US-09-252-991A-6738	Sequence 6738, Ap
35	42.6	1.9	787	US-09-621-976-1878	Sequence 1878, Ap
36	42.6	1.9	2346	US-09-252-991A-6782	Sequence 6782, Ap
37	42.4	1.9	630	US-09-252-991A-3386	Sequence 3386, Ap
38	42.4	1.9	1590	US-09-902-540-9658	Sequence 9658, Ap
39	42.4	1.9	1647	US-09-252-991A-3344	Sequence 3344, Ap
40	42.4	1.9	15923	US-09-902-540-1095	Sequence 1095, Ap
41	42.2	1.9	1089	US-09-796-766-13	Sequence 13, Appl
42	42.2	1.9	1753	US-09-436-699C-15	Sequence 15, Appl
43	42	1.9	804	US-09-252-991A-14038	Sequence 14038, A
44	42	1.9	807	US-09-252-991A-14227	Sequence 14227, A
45	42	1.9	2083	US-09-716-129-41	Sequence 41, Appl

ALIGNMENTS

```
RESULT 1
US-09-949-016-4901
: Sequence 4901, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: FILE REFERENCE: C1001307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: PRIOR FILING DATE: 2000-04-14, 755
: PRIOR APPLICATION NUMBER: 60/241, 755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237, 768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231, 498
: PRIOR FILING DATE: 2000-09-08
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 4901
: LENGTH: 2171
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-4901

Query Match          97.3%; Score 2167.8; DB 4; Length 2171;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2169; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      17 CTAAGTCGCGAGACCTTCTGTGGATCATCGGACCCACTTGGAACCCCACTGACCCAG 76
DB      1 CTAAGTCGCGAGACCTTCTGTGGATCATCGGACCCACTTGGAACCCCACTGACCCAG 60

QY      77 CCCACCTGCTGCAGCCCACTGCGGATGACCATCACTTACCAAGCCAAAGGCGTAA 136
DB      61 CCCACCTGCTGCAGCCCACTGCGGATGACCATCACTTACCAAGCCAAAGGCGTAA 120

QY      137 TGCCCGCTTAGGAGCTCTTCTCCGCTGCTGCTGCTGCGGGGAGACATCTCAAGCT 196
DB      121 TGCCCGCTTAGGAGCTCTTCTCCGCTGCTGCTGCTGCTGCGGGGAGACATCTCAAGCT 180

QY      197 GCTATATGCGAGGTTCTTAATCTTCTGCTGCTGCTGCTACTATCATCCGCTTATTTATAG 256
DB      181 GCTATATGCGAGGTTCTTAATCTTCTGCTGCTGCTGCTACTATCATCCGCTTATTTATAG 240

QY      257 GCTGCGCCCTTACGGAAGAAACAACAGCTGATGTTGAAACAGTACTGTATTGCGCAG 316
DB      241 GCTGCGCCCTTACGGAAGAAACAACAGCTGATGTTGAAACAGTACTGTATTGCGCAG 300

QY      317 CTACATCAGCTCATCCCATTTCTCTGCTGCTGCTGCTTACTACGTACGCTGTGTGAC 376
```

Db 301 CTACATCAGCTATCCCATTTCTTCGTCGTGGCTTCTACGTACCGTGGTGTGAC 360
Qy 377 CCGGTGGTGAACCAAGTACGAGAACTGGCCGTGACCGGCTCATGAGCCTGTGTGC 436
Db 361 CCGGTGGTGAACCAAGTACGAGAACTGGCCGTGACCGGCTCATGAGCCTGTGTGC 420
Qy 437 GGGCTTGTGGAAGGCAAGAGCAAGGCGGGCTGTGGGGGCAAGCTCATCCGCTA 496
Db 421 GGGCTTGTGGAAGGCAAGAGCAAGGCGGGCTGTGGGGGCAAGCTCATCCGCTA 480
Qy 497 CGCCAACTGGGCAAGCTGTCTATCTTGGCAGCGTACAGACCGCAGTACAGGCTT 556
Db 481 CGCCAACTGGGCAAGCTGTCTATCTTGGCAGCGTACAGACCGCAGTACAGGCTT 540
Qy 557 CCGCAGCGCCAGACCTGGTGTGCAAGCAAGCTTTAAGTACCTGGGCAAGCAAGCTT 616
Db 541 CCGCAGCGCCAGACCTGGTGTGCAAGCAAGCTTTAAGTACCTGGGCAAGCAAGCTT 600
Qy 617 GGAAGAACTGAGCTTACCAACAACTGTTCTGGGTCCTGGGTCGTGGTGGTGGCAACT 676
Db 601 GGAAGAACTGAGCTTACCAACAACTGTTCTGGGTCCTGGGTCGTGGTGGTGGCAACT 660
Qy 677 GTCAATGAAGGCGTGGCTTGAAGGTGGAATCCGGGACCTTATCTTCCAGAGCTGTCT 736
Db 661 GTCAATGAAGGCGTGGCTTGAAGGTGGAATCCGGGACCTTATCTTCCAGAGCTGTCT 720
Qy 737 GAACGAGATGAACAACCTTGCATCTAGGTGGAACAACCTGTAATGCTAAGCTGATAG 796
Db 721 GAACGAGATGAACAACCTTGCATCTAGGTGGAACAACCTGTAATGCTAAGCTGATAG 780
Qy 797 TATCCCACTGTGTATACACAGGTGTGTAATGTGGCGGTGTATACAGCTTCTTCTGACTTG 856
Db 781 TATCCCACTGTGTATACACAGGTGTGTAATGTGGCGGTGTATACAGCTTCTTCTGACTTG 840
Qy 857 TCTAGTGGGCGGCGGCTTGTGAACCCAGGCAAGGCTTACCTGGCTTATGACTGACCT 916
Db 841 TCTAGTGGGCGGCGGCTTGTGAACCCAGGCAAGGCTTACCTGGCTTATGACTGACCT 900
Qy 917 CGTTGTGCGCGCTTCAAGCTTCTGCAAGTCTTCTTATGTTGGTGGCTGGAAGTGGC 976
Db 901 CGTTGTGCGCGCTTCAAGCTTCTGCAAGTCTTCTTATGTTGGTGGCTGGAAGTGGC 960
Qy 977 AGAGAGCTCATCAACCCCTTGTGAAGAGATGATGATTTTGAACCACTGGAATGT 1036
Db 961 AGAGAGCTCATCAACCCCTTGTGAAGAGATGATGATTTTGAACCACTGGAATGT 1020
Qy 1037 CGACAGAAATTTGAGAGTGTCCCTGTGTGTGATGAGATGACAGAGACCTGCTTCG 1096
Db 1021 CGACAGAAATTTGAGAGTGTCCCTGTGTGTGATGAGATGACAGAGACCTGCTTCG 1080
Qy 1097 GATGAGACCGGCAATGTAAGTGAATTAAGCCGAGCCCAAGCCCTTACAGCTGCTTC 1156
Db 1081 GATGAGACCGGCAATGTAAGTGAATTAAGCCGAGCCCAAGCCCTTACAGCTGCTTC 1140
Qy 1157 CGCCCAAGTTCGTCGAGCTTCTTATGGGCTTCACTTCAATCACTGCTGAACAAAGA 1216
Db 1141 CGCCCAAGTTCGTCGAGCTTCTTATGGGCTTCACTTCAATCACTGCTGAACAAAGA 1200
Qy 1217 GAGAGTGAAGTTCAGCCCAATCAAGAGAGAGAGAGATGCTCAAGCTGGCATCTTGG 1276
Db 1201 GAGAGTGAAGTTCAGCCCAATCAAGAGAGAGAGAGATGCTCAAGCTGGCATCTTGG 1260
Qy 1277 CCGCTTCTTGAAGGCTGTGATCCCATGATGACATCTTCCAGAGGCAAACTCAAGAGCAA 1336
Db 1261 CCGCTTCTTGAAGGCTGTGATCCCATGATGACATCTTCCAGAGGCAAACTCAAGAGCAA 1320
Qy 1337 ACTACTGTGCGCCAAAGAGGAAATCCCTTTCACAGAGGCTGCGCCAAACCAAGAGC 1396
Db 1321 ACTACTGTGCGCCAAAGAGGAAATCCCTTTCACAGAGGCTGCGCCAAACCAAGAGC 1380
Qy 1397 AGCCAAACAGAACTGTTAGGGGCGAGAGAGCAAAAGGCTGTGAAGCTTGAAGCTGTGGA 1456

Db 1361 AGCCAAACAGAACGTTAGGGGCGAGAGAGCAAAAGGCTGGAAGCTTAAAGCTGTGGA 1440
Qy 1457 CGCCTTCAAGTCTGAGCCCACTGTATCAGAGGCGAGGCTTACTAGAGTCCCAAGAGGCC 1516
Db 1441 CGCCTTCAAGTCTGAGCCCACTGTATCAGAGGCGAGGCTTACTAGAGTCCCAAGAGGCC 1500
Qy 1517 CCTCAGCCCACTCCCATGTTCTTCCCTAGAAACCATCAGCGCCGCTCAAGCTTCAAG 1576
Db 1501 CCTCAGCCCACTCCCATGTTCTTCCCTAGAAACCATCAGCGCGCTCAAGCTTCAAG 1560
Qy 1577 TGTCAAGGCAATACACCAAGAGCAAAAGCTTAAAGACTGTGAGTCTTGGGCGCAAGAA 1636
Db 1561 TGTCAAGGCAATACACCAAGAGCAAAAGCTTAAAGACTGTGAGTCTTGGGCGCAAGAA 1620
Qy 1637 AAGTTTGAATTTGCTCTCAGAGAGCGATGGGGCTTGTATGAGACCAACCAAGATATCTCA 1696
Db 1621 AAGTTTGAATTTGCTCTCAGAGAGCGATGGGGCTTGTATGAGACCAACCAAGATATCTCA 1680
Qy 1697 AGTGAAGAGGAAACTGTGAGTTTAACTGACGATATGCCAGAGATCCCGAAATCA 1756
Db 1681 AGTGAAGAGGAAACTGTGAGTTTAACTGACGATATGCCAGAGATCCCGAAATCA 1740
Qy 1757 CTTCAAGAACTTTGGAAACATCACCAACCAATACATACATCACTCAAGATCAACT 1816
Db 1741 CTTCAAGAACTTTGGAAACATCACCAACCAATACATACATCACTCAAGATCAACT 1800
Qy 1817 GGAATCTTATTTGGGCTTGTGAAAAACAGGATGAGACATTCCTAACCTTCTCTATG 1876
Db 1801 GGAATCTTATTTGGGCTTGTGAAAAACAGGATGAGACATTCCTAACCTTCTCTATG 1860
Qy 1877 GGGATGCTTGGCAGGACAGTCTCTCACTGTGTGTATACACAGAGAGCACTGATCAATG 1936
Db 1861 GGGATGCTTGGCAGGACAGTCTCTCACTGTGTGTATACACAGAGAGCACTGATCAATG 1920
Qy 1937 ACAGCCATACAGTGTCTCACTGTAAGAACTGTCTCAACAGCCTGATCAATGATG 1996
Db 1921 ACAGCCATACAGTGTCTCACTGTAAGAACTGTCTCAACAGCCTGATCAATGATG 1980
Qy 1997 AGCTTAATGATTAATAATCCCACTACCTTCAAGCTTTATGCTTTTATTAATAAAG 2056
Db 1981 AGCTTAATGATTAATAATCCCACTACCTTCAAGCTTTATGCTTTTATTAATAAAG 2040
Qy 2057 TGTGAAGCTTACAGTGAACCATTTGGAACATTTAATCACTGAGCTGGAATCAGAGTCCGG 2116
Db 2041 TGTGAAGCTTACAGTGAACCATTTGGAACATTTAATCACTGAGCTGGAATCAGAGTCCGG 2100
Qy 2117 AACCTTACTTATCTGAATCAAGAGAGCCACACTTATATATCTGCCAAACTAATG 2176
Db 2101 AACCTTACTTATCTGAATCAAGAGAGCCACACTTATATATCTGCCAAACTAATG 2160
Qy 2177 AGTTAATAAA 2187
Db 2161 AGTTAATAAA 2171

RESULT 2
US-09-949-016-16644/c
; Sequence 16644, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

```

; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16644
; LENGTH: 7609
; TYPE: DNA
; ORGANISM: Human
; OS-09-949-016-16644

```

Query Match	28.8%	Score 642.2;	DB 4;	Length 7609;
Best Local Similarity	99.5%	Pred. No. 4,1e-163;		
Matches 644;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0

Query Match	Similarity	99.5%	Pred. No.	4,1e-163;	
Best Local	Conservative	0;	Mismatches	3;	
Matches	644;	Indels	0;	Gaps	0;
Qy	1202	CAGCCTGAACAAAGAGATGATGATTCTCACGCCCAATCAGAGAGACGAGAGATGTCTCA	1267		
Db	7382	CAGCCTGAACAAAGAGAGATGATGATTCTCACGCCCAATCAGAGAGACGAGAGATGTCTCA	7322		
Qy	1262	CGCTGGCATCATTTGGCCGCTTTCCTTAAGGCCTGTGACGTCCCATGATCAACAATCTCTCCAGAGGC	1322		
Db	7322	CGCTGGCATCATTTGGCCGCTTTCCTTAAGGCCTGTGACGTCCCATGATCAACAATCTCTCCAGAGGC	7265		
Qy	1322	AAACTCAAAGAACCAAACTACTGTGGCCCCAAGAGGAAATCCTTCTTCCACGAGGGCCTTGCC	1382		
Db	7262	AAACTCAAAGAACCAAACTACTGTGGCCCCAAGAGGAAATCCTTCTTCCACGAGGGCCTTGCC	7203		
Qy	1382	CAAAAACCAAGAAGGACGCAACAGCTTAAGGGGCCAGAGAACCAACAGGCCTTGGAA	1442		
Db	7202	CAAAAACCAAGAAGGACGCAACAGCTTAAGGGGCCAGAGAACCAACAGGCCTTGGAA	7143		
Qy	1442	GCTTAAGGCTGTGACGCGCTTCAAGTCTTGSCCCATCTGTATCAGAGGCGAGCTACTAAG	1502		
Db	7142	GCTTAAGGCTGTGACGCGCTTCAAGTCTTGSCCCATCTGTATCAGAGGCGAGCTACTAAG	7083		
Qy	1502	TGCCCCACAGACGCGCCCTCAGCGCCCACCTCCAATGTTCTTCCCTTAGAACCATCAGCGCC	1562		
Db	7082	TGCCCCACAGACGCGCCCTCAGCGCCCACCTCCAATGTTCTTCCCTTAGAACCATCAGCGCC	7023		
Qy	1562	GTCAAAGCTTCAACATGTCAACGAGATAGACACCAAAACAAAGACTTAAAGACTGTGAG	1622		
Db	7022	GTCAAAGCTTCAACATGTCAACGAGATAGACACCAAAACAAAGACTTAAAGACTGTGAG	6963		
Qy	1622	TTCTGGGGGCCAAGAAAAAGTTTTGAATTGCTCTCAGAGAGCGATGGGGCCTTGATGTAGACA	1682		
Db	6962	TTCTGGGGGCCAAGAAAAAGTTTTGAATTGCTCTCAGAGAGCGATGGGGCCTTGATGTAGACA	6903		
Qy	1682	CCCAAGAAATATCTCAAGTAGAGAGAGAAACTGTGAGTTTAACTGACGGAATATCCAGA	1742		
Db	6902	CCCAAGAAATATCTCAAGTAGAGAGAGAAACTGTGAGTTTAACTGACGGAATATCCAGA	6843		
Qy	1742	GATCCCCGAAATACTCACTCAAGAAACCTTTGGAAACATCACCACCAACATACACACTAC	1802		
Db	6842	GATCCCCGAAATACTCACTCAAGAAACCTTTGGAAACATCACCACCAACATACACACTAC	6783		
Qy	1802	ACTCAAAAGATCAACATGATCCTTATTTGGGCTTGTGAAAAACAGGAGATG	1848		
Db	6782	ACTCAAAAGATCAACATGATCCTTATTTGGGCTTGTGAAAAACAGGAGATG	6736		

```

RESULT 3
US-09-949-016-16643
? Sequence 16643, Application US/09949016
? Patent No. 6812339
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: C1001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
?

```

```

? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: PARSSEQ for Windows Version 4.0
? SEQ ID NO: 16643
? LENGTH: 18075
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-16643

```

Query Match	28.8%	Score 642.2;	DB 4;	Length 18075;
Best Local Similarity	99.5%;	Pred. No. 7.1e-163;		
Matches 644; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

Query Match	28.8%	Score 642.2	DB 4	Length 18075
Best Local Similarity	99.5%	Pred. No. 7.1e-163		
Matches 644	Conservative 0	Mismatches 3	Indels 0	Gaps 0
QY	1202	CAGCTGAAACAAAGAGAGATGGAGTTCCAGGCCAATCAGAGAGACGAGAGATGTCTCA	1261	
DB	13879	CAGCTTGAAACAAAGAGAGATGGAGTTCCAGGCCAATCAGAGAGACGAGAGATGTCTCA	13938	
OY	1262	CGCTGGCATATATTGGCCGCTTCTTAGGCTCGAGTCCCATATACATTTCTTCCAGGGC	1321	
DB	13939	CGCTGGCATATTGGCCGCTTCTTAGGCTCGAGTCCCATATATCTCTCCAGGGC	13998	
OY	1322	AAACTCAAGACCAAACTACTGTGGGCCCAAGAGGGAATCCCTTCTCAAGAGGCTCTCC	1381	
DB	13999	AAACTCAAGACCAAACTACTGTGGGCCCAAGAGGGAATCCCTTCTCAAGAGGCTCTCC	14058	
OY	1382	CAAAAACACACAGGCGACCCAAACAGACCTTAGGGCCACGAAAGACACAGAGCCTGAA	1441	
DB	14059	CAAAAACACACAGGCGACCCAAACAGACCTTAGGGGCCAGAAAGACACAGAGCCTGAA	14118	
OY	1442	GCATTAGAGCTGTGGAGGCGCTTCAAGCTGTGGCCCATGTATCGAAGGCCAGGCTATAG	1501	
DB	14119	GCTTAGAGCTGTGGAGGCGCTTCAAGCTGTGGCCCATGTATCGAAGGCCAGGCTACTAG	14178	
OY	1502	TGCCCCACAGACGCGCCCTCAGGCCCACTGCCATGTTCTTCCCTTAGAACCATCAGCGCC	1561	
DB	14179	TGCCCCACAGACGCGCCCTCAGGCCCACTGCCATGTTCTTCCCTTAGAACCATCAGCGCC	14238	
OY	1562	GTCAAAGCTTCAAGGTGTACAGGCTATACACCAAAGCAAAACCTTAAAGACTGTAG	1621	
DB	14239	GTCAAAGCTTCAAGGTGTACAGGCTATACACCAAAGCAAAACCTTAAAGACTGTAG	14298	
OY	1622	TTCGTGGGCCCAAGAAAGTTTGAATTGTCTCAGAGAGCGATGGGGCCTTGATGAGCA	1681	
DB	14299	TTCGTGGGCCCAAGAAAGTTTGAATTGTCTCAGAGAGCGATGGGGCCTTGATGAGCA	14358	
OY	1682	CCCAAGATATCTCAAGTGAAGGAGAAACTGTGAGTTTAACTCTGACGATATGCCGA	1741	
DB	14359	CCCAAGATATCTCAAGTGAAGGAGAAACTGTGAGTTTAACTCTGAGGATATGCCGA	14418	
OY	1742	GATCCCGAAATCACTCAAGAAACCTTTGGAAACATCAACCAACATATACACTATAC	1801	
DB	14419	GATCCCGAAATCACTCAAGAAACCTTTGGAAACATCAACCAACATATACACTATAC	14478	
OY	1802	ACTCAAGATCACATGATCTTTATTGGGCTTTGAAAAACAGGATG	1848	
DB	14479	ACTCAAGATCACATGATCTTTATTGGGCTTTGAAAAACAGGATG	14525	

```

RESULT 4
US-09-949-016-4902/c
; Sequence 4902, Application US/09949016
; Patent No. 6812339
;
GENERAL INFORMATION:
;
APPLICANT: VENTER, J. Craig et al.
;
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
;
CURRENT APPLICATION NUMBER: US/09/949,016
;
CURRENT FILING DATE: 2000-04-14
;
PRIOR APPLICATION NUMBER: 60/241,755
;
PRIOR FILING DATE: 2000-10-20
;
PRIOR APPLICATION NUMBER: 60/237,768
;
PRIOR FILING DATE: 2000-10-03
;
PRIOR APPLICATION NUMBER: 60/231,498
;

```

;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 4902
;; LENGTH: 1462
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-4902

Query Match 16.1%; Score 358.4; DB 4; Length 1462;
Best Local Similarity 99.7%; Pred. No. 1.4e-86;
Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1842 AGGATGAGACCACTTCTTAACCTGCTTCTTAATGGGATGCTTCCAGCCAGGCTCTC 1901
DB 1385 AGGATGAGACCACTTCTTAACCTGCTTCTTAATGGGATGCTTCCAGCCAGGCTCTC 1326
QY 1902 ACTGTGTGTACACGAGACGACACTGATCCAGTCCAGCCATACAGCTGTCCACTGA 1961
DB 1325 ACTGTGTGTACACGAGACGACACTGATCCAGTCCAGCCATACAGCTGTCCACTGA 1266
QY 1962 AGAAGTGTCTTACACGCTGATCAATGTTAGTTATGATTAATCCAGAC 2021
DB 1265 AGAAGTGTCTTACACGCTGATCAATGTTAGTTATGATTAATCCAGAC 1206
QY 2022 TACTTCAGCCTTTATGCTTTTATTCATAAACTGGAAGCTGAGACTGAACCTTGG 2081
DB 1205 TACTTCAGCCTTTATGCTTTTATTCATAAACTGGAAGCTGAGACTGAACCTTGG 1146
QY 2082 AATCATTTAACTCACTCTGATTCAGATCGAGAGTGGAAACCTTATGTTCTGATCCAA 2141
DB 1145 AATCATTTAACTCACTCTGATTCAGATCGAGAGTGGAAACCTTATGTTCTGATCCAA 1086
QY 2142 GACAGCCACACCTTAGTATGCTGCCAACTAATGATTTAATTAATCAATCTGTT 2201
DB 1085 GACAGCCACACCTTAGTATGCTGCCAACTAATGATTTAATTAATCAATCTGTT 1026

RESULT 5

US-09-949-016-173056
;; Sequence 173056, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 173056
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-173056

Query Match 8.5%; Score 188.4; DB 4; Length 601;

Best Local Similarity 99.5%; Pred. No. 1e-40;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1842 AGGATGAGACCACTTCTTAACCTGCTTCTTAATGGGATGCTTCCAGCCAGGCTCTC 1901
DB 412 AGGATGAGACCACTTCTTAACCTGCTTCTTAATGGGATGCTTCCAGCCAGGCTCTC 471
QY 1902 ACTGTGTGTACACGAGACGACACTGATCCAGTCCAGCCATACAGCTGTCCACTGA 1961

DB 472 ACTGTGTGTACACGAGACGACACTGATCCAGTCCAGCCATACAGCTGTCCACTGA 531
QY 1962 AGAAGTGTCTTACACAGCCTGATCAATGTTAGTTATGATTAATCCAGAC 2021
DB 532 AGAAGTGTCTTACACAGCCTGATCAATGTTAGTTATGATTAATCCAGAC 591
QY 2022 TACTTCAGCC 2031
DB 592 TACTTCAGCC 601

RESULT 6

US-09-949-016-173051
;; Sequence 173051, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 173051
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-173051

Query Match 7.2%; Score 161.4; DB 4; Length 601;
Best Local Similarity 99.4%; Pred. No. 2.1e-33;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 582 GCAAGCTTTATGACTCCGGGAGACACAGCACTTGGAGAAATGAGCTTACACACAC 641
DB 321 GCAAGCTTTATGACTCCGGGAGACACAGCACTTGGAGAAATGAGCTTACACACAC 380
QY 642 ATGTTCCGGTGGCTCGGGGTTGTTTCCCACTGTCATGAAGGCTGTTGAGGT 701
DB 381 ATGTTCCGGTGGCTCGGGGTTGTTTCCCACTGTCATGAAGGCTGTTGAGGT 440
QY 702 CGAATCCGGGACCTTATCTGCTCCAGAGCTGCTGAACGAGA 744
DB 441 CGAATCCGGGACCTTATCTGCTCCAGAGCTGCTGAACGAGA 483

RESULT 7

US-09-949-016-173052
;; Sequence 173052, Application US/09949016
;; Patent No. 6812339
;; GENERAL INFORMATION:
;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: CL001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 173052

LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-173052

Query Match 7.2%; Score 161.4; DB 4; Length 601;
Best Local Similarity 99.4%; Pred. No. 2.1e-33;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 582 GCAGGCTTTATGACTCCGGCAGAACACAGAGATTGGAAACTGAGCTTACACACAC 641
DB 30 GCAGGCTTTATGACTCCGGCAGAACACAGAGATTGGAAACTGAGCTTACACACAC 89
QY 642 ATGTTCTGGGTGCTCCGCTGGGTGGTGGTCCCAACCTGTAATGAAGGCGGTGGCTTGAAGT 701
DB 90 ATGTTCTGGGTGCTCCGCTGGGTGGTGGTGGTCCCAACCTGTAATGAAGGCGGTGGCTTGAAGT 149
QY 702 CGAATCCGGGACCCCTATCTCTCCAGAGAGCTGCTGAACGAGA 744
DB 150 CGAATCCGGGACCCCTATCTCTCCAGAGAGCTGCTGAACGAGA 192

RESULT 8
US-09-949-016-173045

Sequence 173045, Application US/0949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 173045

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-173045

Query Match 5.7%; Score 128; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.5e-24;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 GACCCAAAGCCCACTGCTGACAGCCCACTGCTGGCATGACCATCACTTACAAAGCCAA 128
DB 474 GACCCAAAGCCCACTGCTGACAGCCCACTGCTGGCATGACCATCACTTACAAAGCCAA 533
QY 129 GTGCTATATGCGCGCTTGGCTCTTCCCGCTGCTGCTGTGTGCGGGGAGCATC 188
DB 534 GTGCTATATGCGCGCTTGGCTCTTCCCGCTGCTGCTGTGTGCGGGGAGCATC 593
QY 189 TACAAGCT 196
DB 594 TACAAGCT 601

RESULT 9
US-09-270-767-735

Sequence 735, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 735
LENGTH: 419
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-735

Query Match 4.6%; Score 102.4; DB 4; Length 419;
Best Local Similarity 62.9%; Pred. No. 1.7e-17;
Matches 176; Conservative 0; Mismatches 101; Indels 3; Gaps 1;

QY 944 GTTCTTCTTCAATGTTGGCTGGCTGGAAGGTGGAGAGAGCTCATCAACCCCTTTGAGA 1003
DB 1 GTTCTTCTTCAATGTTGGCTGGCTGGAAGGTGGAGAGAGCTCATCAACCCCTTTGAGA 60
QY 1004 GGATGATGATGATTTTGAAGCAACCTGATTTGACAGAAATTTGACAGGTGCTGTT 1063
DB 61 AGACGATGATGATTTTGAAGCAACCTGATTTGACAGAAATTTGACAGGTGCTGTT 120
QY 1064 GCGTGGATGATGATGACCAAGACCTGCTGGATGAGCCGACATGATCTGAATTA 1123
DB 121 GATGCTGACAGAGATGACCAAGACCTGCTGGATGAGCCGACATGATCTGAATTA 180
QY 1124 GCCCGAGCC--ACAGCCCCCTACACAGCTGCTTCCGCCAGTTCCGTCGAGCCTCTT 1180
DB 181 GGTGTTCCCAACGAGCTGCCCTTACACATGCTGCGAATGATTTCCGGAGAAATCATCC 240
QY 1181 TATGGGCTCCACTTCAACATCAAGCTGAAACAAAGAGAG 1220
DB 241 AGAGCGTCCACTGCAAGATCGAGTGGCCCAAGAAATGCG 280

RESULT 10
US-09-270-767-16017

Sequence 16017, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 16017

LENGTH: 419

TYPE: DNA

ORGANISM: Drosophila melanogaster

US-09-270-767-16017

Query Match 4.6%; Score 102.4; DB 4; Length 419;
Best Local Similarity 62.9%; Pred. No. 1.7e-17;
Matches 176; Conservative 0; Mismatches 101; Indels 3; Gaps 1;

QY 944 GTTCTTCTTCAATGTTGGCTGGCTGGAAGGTGGAGAGAGCTCATCAACCCCTTTGAGA 1003
DB 1 GTTCTTCTTCAATGTTGGCTGGCTGGAAGGTGGAGAGAGCTCATCAACCCCTTTGAGA 60
QY 1004 GGATGATGATGATTTTGAAGCAACCTGATTTGACAGAAATTTGACAGGTGCTGTT 1063
DB 61 AGACGATGATGATTTTGAAGCAACCTGATTTGACAGAAATTTGACAGGTGCTGTT 120
QY 1064 GCGTGGATGATGATGACCAAGACCTGCTGGATGAGCCGACATGATCTGAATTA 1123
DB 121 GATGCTGACAGAGATGACCAAGACCTGCTGGATGAGCCGACATGATCTGAATTA 180
QY 1124 GCCCGAGCC--ACAGCCCCCTTACACAGCTGCTTCCGCCAGTTCCGTCGAGCCTCTT 1180
DB 181 GGTGTTCCCAACGAGCTGCCCTTACACATGCTGCGAATGATTTCCGGAGAAATCATCC 240
QY 1181 TATGGGCTCCACTTCAACATCAAGCTGAAACAAAGAGAG 1220

Db 241 AGAGCGCTCCACTGCCAAGATCGAGTGGCCCAAGATGG 280

RESULT 11

US-09-949-016-173053
; Sequence 173053, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-03

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 173053

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-173053

Query Match
Best Local Similarity 3.2%; Score 71.4; DB 4; Length 601;
Matches 75; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 756 CGTACTAGTGTGACACCTGTATGCTTACGACTGGATTAGTCCACTGGTGTATACA 815
Db 1 CGTACTAGTGTGACACCTGTATGCTTACGACTGGATTAGTCCACTGGTGTATACA 60

Qy 816 CAGGTGTGCTGCTGCGGTG 836
Db 61 CAGGTGTGCTGCTGCGGTG 81

RESULT 12

US-09-949-016-173054
; Sequence 173054, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 173054

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-173054

Query Match
Best Local Similarity 3.2%; Score 70.4; DB 4; Length 601;
Matches 74; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 757 GTACTAGTGTGACACCTGTATGCTTACGACTGGATTAGTCCACTGGTGTATACA 816
Db 61 GTACTAGTGTGACACCTGTATGCTTACGACTGGATTAGTCCACTGGTGTATACA 60

Db 1 GTACTAGTGTGACACCTGTATGCTTACGACTGGATTAGTCCACTGGTGTATACA 60
Qy 817 AGGTGTGCTGCTGCGGTG 836
Db 61 AGGTGTGCTGCTGCGGTG 80

RESULT 13

US-09-949-016-173053/c
; Sequence 173053, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-03

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 173053

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-173053

Query Match
Best Local Similarity 3.0%; Score 66; DB 4; Length 601;
Matches 78; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1864 CTGCTTCTTATGAGGATCTTCGCGAGCGTCTTACCTGTGTGTACACGACGAGA 1923
Db 103 CAGTTTCCCAAAAGGCGACCTTACGAGCTTATGCTTCTTATGAGGATCTTCGCGAGCGTCTTACCTGTGTGTACACGAGTGA 44

Qy 1924 CACTGATCCAGTCAACGCCATACAGCTGTCCACTGA 1961
Db 43 TACTATCCAGTCTGAGGATATACAGTGTCCACTGA 6

RESULT 14

US-09-949-016-173054/c
; Sequence 173054, Application US/09949016
; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 173054

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-173054

Query Match
Best Local Similarity 3.0%; Score 66; DB 4; Length 601;
Matches 78; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Oy 1864 CTGCTTCTTAATGGGAGATGCTTCGCGACAGAGTCTCAACCGTGTGTAAACACAGCAGGA 1922
 Db 102 CAGTTTCCCAAAAAGGCGAGCCCAACCACTAGTCTCACTGTGTATTACACAGTGGGA 43
 Oy 1924 CACTGATTCAGTCACACGCCATTACAGCTGTCCACACTGA 1961
 Db 42 TACTATATCCAGTCGTAGGCATTACAGATGTCCACACTGA 5

RESULT 15
US-09-949

US-09-949-016, 16643/c
 Sequence 16643, Application US/09949016
 Patent No. 6812339
 GENERAL INFORMATION:
 APPLICANT: VENTER, J. Craig et al.
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 FILE REFERENCE: CL001307
 CURRENT APPLICATION NUMBER: US/09/949,016
 CURRENT FILING DATE: 2000-04-14
 PRIOR APPLICATION NUMBER: 60/241,755
 PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: ParSeq for Windows Version 4.0
 SEQ ID NO 16643
 LENGTH: 18075
 TYPE: DNA
 ORGANISM: Human
 US-09-949-016-16643

Query Match	3.0%;	Score 66;	DB 4;	Length 18075;
Best Local Similarity	79.6%;	Pred. No. 1.3e-06;		
Matches	78;	Conservative	0;	Mismatches 20;
			Indels	0;
			Gaps	0;

QY	1864	CTGCTTCTTAATGGGAGTGGCTTGGCCAGCCAGGCTCTCACTGTGTGACACAGCAGAGA	12933
Db	9126	CAGTTTCCCAAAAAGGCGAGCCCAACAGCCAGTGTCTCACTGTGTATATACACAGTGGGA	9065
QY	12924	CACGTGATCTCAAGTCAACGCCATTAACAGCTGTCCCACTGTA	1361
Db	9066	TACTAATCTCAGTCTGTAGGCATTAACAGTGTCCACACTGTA	9029

Search completed: March 26, 2005, 13:49:34
Job time : 357.338 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 23, 2005, 06:37:09 ; Search time 7468.71 Seconds
(without alignments)
3795.340 Million cell updates/sec

Title: US-09-622-964A-3
Perfect score: 3120
Sequence: 1 MTTTYSQVNNARLGSFSRL.....TLKDHMDPYALENREDAHS 565

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09622964/rnatc_22032005_062218_26369/arp_query.fasta_1.2069
-DB=GenEmbl -OPMT=faastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blowsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09622964.@CGN_1.1_15531.@rnatc_22032005_062218_26369 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database : GenEmbl.*
1: gb_ba:*
2: gb_hc:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	3116	99.9	2229	6 BD136720 Best's ma
2	3110	99.7	1758	6 AX745964 Sequence
3	3110	99.7	1758	9 AY515704 Homo sapi
4	3110	99.7	2171	9 AF073501 Homo sapi

5	3110	99.7	2210	9 AF057169	AF057169 Homo sapi
6	3083	98.8	2170	6 CQ729633	CQ729633 Sequence
7	3047.5	97.7	2429	6 BD136721	BD136721 Best's ma
8	3041.5	97.5	2420	9 AF057170	AF057170 Homo sapi
9	3001	96.2	2187	9 AY357925	AY357925 Macaca fa
10	2817	90.3	2435	9 BC041664	BC041664 Homo sapi
11	2672.5	85.7	2441	9 BC015220	BC015220 Homo sapi
12	1914	61.3	2035	10 BC079048	BC079048 Rattus no
13	1899	60.9	1916	6 BD136744	BD136744 Best's ma
14	1896	60.8	1904	10 AY450427	AY450427 Mus muscu
15	1701	54.5	1289	4 AY064707	AY064707 Sus bcto
16	1655.5	53.1	112309	4 AC003025	AC003025 Human Chr
17	1652.5	53.0	142092	9 AF139813	AF139813 Homo sapi
18	1652.5	53.0	196080	9 AC004228	AC004228 Homo sapi
19	1645.5	52.7	163024	9 AP006260	AP006260 Homo sapi
20	1645.5	52.7	163915	2 AC087451	AC087451 Homo sapi
21	1645.5	52.7	166867	2 AP003733	AP003733 Homo sapi
22	1636.5	52.5	133683	2 AC084857	AC084857 Homo sapi
23	1615	51.8	16125	6 BD136719	BD136719 Best's ma
24	1571	50.4	1263	6 BD132659	BD132659 Secreter
25	1433	45.9	160169	2 AC051664	AC051664 Homo sapi
26	1358.5	43.5	1898	5 BC043854	BC043854 Xenopus l
27	1358.5	43.5	1921	5 AY273825	AY273825 Xenopus l
28	1348	43.2	1882	5 CR760914	CR760914 Xenopus t
29	1348	43.2	2021	5 BC061379	BC061379 Xenopus t
30	1343	43.0	2719	10 AY450426	AY450426 Mus muscu
31	1337.5	42.9	2064	5 BC084229	BC084229 Xenopus l
32	1332	42.7	2007	6 AY515706	AY515706 Homo sapi
33	1332	42.7	2500	6 AX834175	AX834175 Sequence
34	1332	42.7	2500	9 AK096459	AK096459 Homo sapi
35	1329.5	42.6	1650	5 AY273826	AY273826 Xenopus l
36	1311.5	42.0	2028	6 AY745968	AY745968 Sequence
37	1296	41.5	1506	9 AF440758	AF440758 Homo sapi
38	1286.5	41.2	1957	10 BC031186	BC031186 Mus muscu
39	1286.5	41.2	1957	10 BC019528	BC019528 Mus muscu
40	1286.5	41.2	1992	10 BC036157	BC036157 Mus muscu
41	1286.5	41.2	2028	10 BC036163	BC036163 Mus muscu
42	1285.5	41.2	1954	10 AY450428	AY450428 Mus muscu
43	1284	41.2	1908	9 AF440756	AF440756 Homo sapi
44	1282	41.1	2137	6 AX745966	AX745966 Sequence
45	1281.5	41.1	1530	6 CQ731444	CQ731444 Sequence

ALIGNMENTS

RESULT 1	BD136720	2229 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	Best's macular dystrophy gene.				
DEFINITION	BD136720				
ACCESSION	BD136720.1 GI:23231665				
VERSION	JP 2002504559-A/2.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 2229)				
AUTHORS	Petrushkin,K., Caskey,T.C., Metzker,M. and Wadelius,C.				
TITLE	Best's macular dystrophy gene				
JOURNAL	Patent: JP 2002504559-A 2 12-FEB-2002;				
COMMENT	MERCK & CO INC, CLABS WADELIIUS				
OS	Homo sapiens (human)				
PN	JP 2002504559-A/2				
PD	12-FEB-2002				
PR	22-FEB-1999 JP 2000533447				
PC	25-FEB-1998 US 60/075941,18-DEC-1998 US 60/112936 PI				
PC	KONSTANTIN PETRUSHKIN,THOMAS C CASKEY,MICHAEL METZKER,CLABS PI				
PC	C07K16/18,C07K14/47,C12N5/10,C12N15/09,C12P19/34,C12Q1/68//				
PC	C12P21/08,				
PC	C12N5/00,C12N15/00				
PC	Best's macular dystrophy gene				
PC	Best's macular dystrophy gene				
PC	Key				
PC	Location/Qualifiers				

FT source 1..2229
FT /organism='Homo sapiens (human)'.
FEATURES
source location/Qualifiers
1..2229
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6 74e-267 Length: 2229
Score: 3116.00 Matches: 584
Percent Similarity: 99.83% Conservative: 0
Best Local Similarity: 99.83% Mismatches: 1
Query Match: 99.87% Indels: 0
DB: Gaps: 0
US-09-622-964a-3 (1-585) x BD136720 (1-2229)
QY 1 MetThrIleThrTyThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
DB 105 ATGACCATCATCTTACACAGCCAGTGGCTAATGCCCGTTAGGCTCTTCCCGCCTG 164
QY 21 LeuLeuCyThrArgGlySerIleTyTyrIleLeuTyGlyLysPheLeuIlePheLeu 40
DB 165 CTGCTGTCTGGCGGGGAGCATCTACAGCTGTATATGCGAGTTCTTAATCTTCCCTG 224
QY 41 LeuCyTyTyTyrIleIleArgPheIleTyTyrArgLeuAlaLeuThiGluGluGlnLeu 60
DB 225 CTCTGTACTATCATATCCGCTTATTATTTATAGCTGGCCCTCAGAGAAACACAGCTG 284
QY 61 MetPheGluLysLeuThrLeuTyTyrCyAspSerTyTyrIleGlnLeuIleProIleSerPhe 80
DB 285 ATGTTTGAAACCTGACTCTGTATTGGCAGCATTCACCTCAGCTCATCCCATTTCCCTTC 344
QY 81 ValLeuGlyPheTyTyrValThrLeuValValThrArgTyrTyrPheGlnTyrGluAsnLeu 100
DB 345 GTGCTGGGCTTTAGCTGAAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 404
QY 101 ProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGln 120
DB 405 CGGTGGCCGACCGGCTCATGAGCTGTGGTGGCTTCGTGAAAGCAAGACAGAGCA 464
QY 121 SerArgLeuLeuArgArgThrLeuIleArgTyTyrAlaAsnLeuGlyAsnValLeuIleLeu 140
DB 465 GAGCGGCTGTGGCGGAGCATCTCATCCGTAACGCAACCTGGGCAACGTCTCATCCTG 524
QY 141 ArgSerValSerThrAlaValTyrTyrValArgPheProSerAlaGlnIleLeuValGlnAla 160
DB 525 CGAGCGTCACGACCGGAGCTTCAAGGCTTCCCGAGCGCCAGCACTGGTGCAAGCA 584
QY 161 GlyPheMetThrProAlaGlnIleIleArgIleLeuGluLysLeuSerLeuProIleAsnMet 180
DB 585 GCGTTTATGACTCCGGCAGAACACAGAGCATTTGAGAACTGACCTCAACCAACAATG 644
QY 181 PheTyrValProTyrValTyrPheAlaAsnLeuSerMetLeuValAlaTyrLeuGlyGlyArg 200
DB 645 TTCTGGGTGCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 704
QY 201 IleArgAspProIleLeuLeuGlnSerIleLeuAsnGluMetAsnThrLeuArgThrGln 220
DB 705 ATCCGGGACCTATCCGTCTCCAGAGCTGTGACAGATGAACACCTTGGCGTAAGTCA 764
QY 221 CysGlyIleLeuTyTyrAlaTyTyrAspTyrIleSerIleProLeuValTyrThrGlnVal 240
DB 765 TGTGGACACCTGTATGCTTCAAGTGAATAGTATCCACCTGGGTGTATACACAGGTG 824
QY 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValAlaGlyArgGlnPheLeuAsnPro 260
DB 825 ACTGTGGCGGTGACAGCTTCTTCTTCACTGTGTCTTGTGGTGGTGGTGGTGGTGGTGGT 884
QY 261 AlaIysAlaTyTyrProGlyIleGluLeuAspLeuValValProValPheThrPheLeuGln 280

DB 885 GCCAAGGCTTACCTGGCGATGAGCTGGAACCTCGTTGGCCGCTTCAAGCTTCTCAG 944
QY 281 PhePhePheTyTyrValGlyTyrLeuLysValAlaGluGlnLeuIleAsnProPheGlyGlu 300
DB 945 TTCTTCTTCTTATGTTGGCTGGCTGAGAGTGGCAGAGAGCTCATCAACCTTTGGAGAG 1004
QY 301 AspAspAspAspPheGluThrAsnTyrIleValAlaAspArgAsnLeuGlnValSerLeu 320
DB 1005 GATGATGATGATTTTGAACCACTGGAATGTGCACAGGAATTTGACAGGTGCTCCTGTTG 1064
QY 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyTyrPheAsnLys 340
DB 1065 GCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1124
QY 341 ProGluProGlnProProTyTyrThrAlaIleSerAlaGlnPheArgArgAlaSerPheMet 360
DB 1125 CCCAGGCCACAGCCCCCTTACAGAGCTGTCCGCCCAAGTTCCGTGAGGCTCCTTATG 1184
QY 361 GlySerThrPheAsnIleSerLeuAsnLysGluGluMetGluPheGlnProAsnGlnGlu 380
DB 1185 GGTTCACCTTCAACATACAGCTGGAACAGAGAGATGAGTTCCAGCCCAATCAGAGAG 1244
QY 381 AspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
DB 1245 GACGAGGAGGATGCTACGCTGGCATCATTTGGCGCTTCTTAGGCTTCGAGTCCATGAT 1304
QY 401 HisHisProProArgAlaAsnSerArgThrLysLeuLeuTyrProLysArgGluSerLeu 420
DB 1305 CACCATCTTCCAGGGGCAACCTCAGAGACCAACCTAGTGGCCCAAGAGGAAATCCCTT 1364
QY 421 LeuHisGluGlyLeuProLysAsnHisLysAlaAlaLysGluAsnValArgGlyGlnGlu 440
DB 1365 CTCACAGAGGCTTGGCCCAAAAACCAAGGCGAGCCAAACAGAACTTAGGGGCCAGAA 1424
QY 441 AspAsnLysAlaTyrLysLeuLysAlaValAlaPheLysSerGlyProLeuTyTyrGln 460
DB 1425 GACAAACAGGCTCGAAGCTTAAAGCTGTGAAGCTTCAAGTCTGACCATGTATCAG 1484
QY 461 ArgProGlyTyTyrSerIleAspProGlnThrProLeuSerProThrProMetPhePhePro 480
DB 1485 AGGCGAGGCTACTACAGTCCCAAGAGAGCCCTCAGGCCCTCCCAATGTTCTTCCCC 1544
QY 481 LeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLys 500
DB 1545 CTAGAACATACAGCGCGCTCAAAAGCTTCAAGTGTCAAGGATGACACCAAAACAA 1604
QY 501 SerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLeuSerGluSerAsp 520
DB 1605 AGCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCAT 1664
QY 521 GlyAlaLeuMetGluHisProGluValSerGlnValArgArgLysThrValGluPheAsn 540
DB 1665 GGGGCTTATGAGAGCAACCAAGATATCTCAAGTGAAGAGAAACGTGTGAAGTTTAC 1724
QY 541 LeuThrAspMetProGluLysProGluAsnHisLeuLysGluProLeuGluGlnSerPro 560
DB 1725 CTGACGAGATATCCAGAGATCCCGAAAATCACTCAAGAAACCTTTGGAACATCA 1784
QY 561 ThrAsnIleHisThrThrLeuLysAspHisMetAspProTyTyrPheAlaLeuGluAsnArg 580
DB 1785 ACCCAACATACACATCACTCAAAAGATCAATGATCTTATTTGGGCTTGGAAAAACAG 1844
QY 581 AspGluAlaHisSer 585
DB 1845 GATGAAGCAATTCC 1859
RESULT 2
AX745964 1758 bp DNA linear PAT 14-MAY-2003
LOCUS AX745964
DEFINITION Sequence 1 from Patent WO03030922.
ACCESSION AX745964
VERSION AX745964.1 GI:30724619
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Steuernagel, A., Bruenner, G., Fritsch, R., Eulenberg, K. and
Clausen, T.
TITLE Bestrophin and bestrophin homologous proteins involved in the
regulation of energy homeostasis
JOURNAL Patent: WO 03030922-A 1 17-APR-2003;
Developmental Biology of the Eye
FEATURES
source location/Qualifiers
1. 1758
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Prod. No.: 1.68e-266 Length: 1758
Score: 3110.00 Matches: 583
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 2
Query Match: 99.68% Indels: 0
Gaps: 0
DB: 6
US-09-622-964A-3 (1-585) x AX745964 (1-1758)
Qy 1 MetThrIleThrYrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db 1 ATGACATCACTTACACAGCAAGTGGCTAATGCCGCTTACGCTCTTCCGGCTG 60
Qy 21 LeuLeuCystrPArgGlySerIleYrIlyLeuLeuYrGlyGluPheLeuIlePheLeu 40
Db 61 CTGCTGCTGCGCGGGCAGCATCTACAAAGCTGCTAATGCGAGTCTTAACTCTTCG 120
Qy 41 LeuCystrYrYrIleIleArgPheIleYrArgLeuAlaLeuThrGluGlnGluLeu 60
Db 121 CTCTGCTACTACATCATCCGCTTATTTATGAGCTGGCCCTCACGAAAGAACACAGCTG 180
Qy 61 MetPheGluYLeuThrIleYrCyAspSerYrIleGlnLeuIleProIleSerPhe 80
Db 181 ATGTTTGAAGAACTGACTCTGTATTTGGCAGCTCAATCCAGCTCATTCCTTC 240
Qy 81 ValLeuGlyPheYrValThrLeuValValThrArgTrpPheAsnGlnYrGluAsnLeu 100
Db 241 GTGCTGGGCTTCTACGTGACGCTGCTGCTGACCCGCTGGTGAACAGTACGAACCTG 300
Qy 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyValAspGluGln 120
Db 301 CCGTGGCCCAACCCCTCATGAGCTGTGTGGGCTTCTCGAAGGCGAAGGACGACCA 360
Qy 121 SerArgLeuLeuArgArgThrLeuIleArgYrAlaAsnLeuGlyAsnValLeuIleLeu 140
Db 361 GGGCGGCTGCGGGCGGCGGCTCATCCGCTACCCCAACCTGGGCAAGTGTCTATCTG 420
Qy 141 ArgSerValSerThrAlaValYrIlyArgPheProSerAlaGlnIleAlaValGlnAla 160
Db 421 CGCAGCGTCAGCAGCCGAGCTCAAGGCTTCCCGAGCGCCAGCACTGGTCAAGCA 480
Qy 161 GlyPheMetThrProAlaGlnIleValGlnLeuGluYLeuSerLeuProIleAsnMet 180
Db 481 GGGCTTTATGACTCGGCGAAGACACAAAGATTGGAAGAACTGAACCTTAACCAACAATG 540
Qy 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLeuValATrPLeuGlyValArg 200
Db 541 TTCTGGGTGCGCTGGGTGTGTGTGCTCAACTGTCAATGAAGGCTGGCTTGAAGTCA 600
Qy 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgThrGln 220
Db 601 ATCCGGAGACCTATCTGCTCCAGAGCTGCTGAACGATGAACACCTTGGCTATCTAG 660

Qy 221 CysGlyHisLeuYrAlaYrAspTrpIleSerIleProLeuValYrThrGlnValVal 240
Db 661 TGTGGACACCTGATATGCTTACAGCTGATATAGATCCACCTGGGTGTATCACAGGTGTG 720
Qy 241 ThrValAlaValYrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
Db 721 ACTGTGGCGGTGTACACCTTCTTCTGACTGTGTCTAGTTGGCGCGAGTTCTCAACCCA 780
Qy 261 AluValAlaYrProGlyHisGluLeuAspLeuValAlaProValPheThrPheLeuGln 280
Db 781 GCCAAGGCTTACCTGGCCATGAGCTGACTCTGTGTGGCTTCTTCACTCTCTGAG 840
Qy 281 PhePhePheYrValGlyTrpLeuValAlaGluGlnLeuIleAsnProPheGlyGlu 300
Db 841 TTCTTCTTCTATGTGGCTGGCTGTGAAGTGGCAGAGCAGCTCATCAACCTTTGGAGAG 900
Qy 301 AspAspAspAspPheGluThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeu 320
Db 901 GATGATGATGATTTTGAACCAACTGATTTGTGACAGAAATTTGACAGTGTCCCTGTG 960
Qy 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetYrTrpAsnIly 340
Db 961 GCTGTGATGAGATGACACAGAGCTGCTCGATGGAGCCGACATGTACTGGAATAG 1020
Qy 341 ProGluProGluProProYrTrpThrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
Db 1021 CCGAGCCACAGCCCTTACACAGCTGTCTCCGCCAGTTCCGTCAAGCTCTTTATG 1080
Qy 361 GlySerThrPheAsnIleSerLeuAsnIlyGluMetGluPheGlnProAsnGlnGlu 380
Db 1081 GGCTCCACCTTCAACATCACTGATCAAGAGAGAGATGAGTCCAGCCCAATCAGAG 1140
Qy 381 AspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
Db 1141 GACGAGAGAGATGCTCAGCTGGCATATTGGCGCTTCTTGAAGCTGCAAGTCCCATGAT 1200
Qy 401 HisIleProProAspAlaAsnSerArgThrIlyLeuLeuTrpProIlyAspGluSerLeu 420
Db 1201 CACCATCTCTCCAGGCGAATCTCAAGSACCAACTACTGTGGCCCAAGAGGAATCCCTT 1260
Qy 421 LeuHisGluGlyLeuProIlyAsnHisIlyValAlaIlyGlnAsnValArgIlyGlnGlu 440
Db 1261 CTCACAGAGGCTGCCCAAAACCAACAGGACAGCAAAAGAACTTGAAGGCGCACAGAA 1320
Qy 441 AspAsnIlyAlaTrpIlyLeuValValAlaAspAlaPheIlySerGlyProLeuYrGln 460
Db 1321 GACAAACAAGGCTGGAACTTAAAGCTGTGAGCCCTTCAAGTCTGCCCACTGTATCTAG 1380
Qy 461 ArgProGlyYrYrSerAlaProGlnThrProLeuSerProThrPrometPhePhePro 480
Db 1381 AGCCAGGCTACTACAGTGGCCCAAGACACCTCTCAAGCCCACTCCCATGTCTTCCC 1440
Qy 481 LeuGluProSerAlaProSerIlyValHisIleSerValThrGlyIleAspThrIlyAspIly 500
Db 1441 CTGAACATCAAGGCGCTCAAGACTTCAAGTGTCAAGGATATGACACCAAGACAA 1500
Qy 501 SerLeuIlyThrValSerSerGlyAlaIlyIlySerPheGluLeuLeuSerGluSerAsp 520
Db 1501 AGCTTAAAGACTGAGTCTTGGGGCCAAAGAAAGTTTGAATTTGCTCTCAAGAGGAT 1560
Qy 521 GlyAlaLeuMetGluHisProGluValIleSerGlnValIleArgArgIlyValGluPheAsn 540
Db 1561 GGGGCTTGTATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAACTGTGAGATTATAC 1620
Qy 541 LeuThrAspMetProGluIleProGluAsnHisIleuIlyGluProLeuGluGlnSerPro 560
Db 1621 CTGACGATATGACAGAGATCCCGAATAATCACTCAAGAACTTTGAAATATCA 1680
Qy 561 ThrAsnIleHisThrThrIleuIlyAspHisIleMetAspProYrTrpAlaLeuGluAsnArg 580
Db 1681 ACCAACTATCACTTACACTCAACAAAGATCAATGATCTTAATGGCTTGGAAAAACAG 1740
Qy 581 AspGluAlaHisSer 585

Db 1741 GATGAGACATCTCC 1755
|||||
RESULT 3
LOCUS AY515704 1758 bp mRNA linear PRI 28-JAN-2004
DEFINITION Homo sapiens bectrophin 1 mRNA, complete cds.
ACCESSION AY515704
VERSION AY515704.1 GI:41216872
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Sun,H., Tsunenari,T., Yau,K.W. and Nathans,J.
TITLE The vitelliform macular dystrophy protein defines a new family of chloride channels
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (6), 4008-4013 (2002)
MEDLINE 21902086
PUBMED 11904445
REFERENCE 2 (bases 1 to 1758)
AUTHORS Tsunenari,T., Sun,H., Williams,J., Cahill,H., Smallwood,P., Yau,K.W. and Nathans,J.
TITLE Structure-function analysis of the bectrophin family of anion channels
J. Biol. Chem. 278 (42), 41114-41125 (2003)
MEDLINE 22917481
PUBMED 12907679
REFERENCE 3 (bases 1 to 1758)
AUTHORS Sun,H., Tsunenari,T., Yau,K.-W. and Nathans,J.
TITLE Direct Submission
Submitted (02-JAN-2004) Molecular Biology, Neuroscience, Johns Hopkins University, 725 N. Wolfe Street, Baltimore, MD 21205, USA
FEATURES
source Location/Qualifiers
1..1758
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1758
/note="chloride channel"
/codon_start=1
/product="bectrophin 1"
/protein_id="AA89654.1"
/db_xref="GI:41216873"
/translation="MTITTSQVANNALGFSRLLCWRGSIYKLYGFLPIILCYI
IIRTYRLALFEQQLMPEKTLTCDSTYIQILPISFVGLVTVTRMNQYENLPW
PDRILMSLVGFVEGKDEQGRILRRTLLIRYANLGNVLIRSVSTAIVYRFPASQHLVOA
GFWTPAEHKQLEKSLPHNMFWVWVFMILSMKAMLGRIIDPILLOSILNEMTLIR
TCGGHYAYDWISLIPYTOVVVAVYSPFLTCVAGROPINAKAYAGHELDLVVPVF
TPIQFFPYGMILKABOLINPGEDEDDPFTNMIVDRNLQVSLAIVDEMDQDLPMBP
DMWNKREPPPTTAAQAQRRASFMGSTNISLNKEMEFQRODEBEBAHAGIICR
FLGQSDHHPRRANSRTKLMPRESLHFGLPKNHAKAKQNVQGEDNKAMKLVKAV
DARSAFLYGRPGYSAPQTPLSPTMPFPLPSAPSKLASHVGIIDKDKSLKTVSSG
AKKSFELISDSDGALMEHPVSQVRRKTFEFLNLDMPBEIPENHLKEPLDSPTNIHTT
LKDHMDPYMALENRDEAHS"
ORIGIN
Alignment Scores:
Pred. No.: 1.68e-266 Length: 1758
Score: 3110.00 Matches: 583
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 2
Query Match: 99.68% Indels: 0
Gaps: 0
US-09-622-964a-3 (1-585) x AY515704 (1-1758)
QY 1 MetThrIleThrTrpThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db 1 ATGACCAATCACTTACACCAAGCGCAAGTGGCTTAATGCCCGCTTAGAGCTCTTCTCCCGGCTG 60
QY 21 LeuLeuCystrpArgGlySerIleTrpIleLeuLeuTrpGlyGluPheLeuIlePheLeu 40

Db 61 CTGCTGATGTCGGCGGACACATCTACAGCTGATATGACGAGTCTTAATCTTCCTG 120
QY 41 LeuCystrpTrpIleIleIleArgPheIleTrpArgLeuAlaLeuThrGlyGluGlnGlnLeu 60
Db 121 CTCTGCTACATCAATCAACCCCTTATTATTAAGCTGGCCCTCACGGAAGAACAACACTG 180
QY 61 MetPheGluTrpLeuThrLeuTrpCysArgSerTrpIleGlnLeuIlePheIleSerPhe 80
Db 181 ATGTTTAGAAATGACTCTGTATTGGAGACAGCTACACATCCAGCTCATCCCATTTCTTC 240
QY 81 ValLeuGlyPheTrpValThrLeuValValThrArgTrpTrpAsnGlnTrpGluAsnLeu 100
Db 241 GTGCTGGGCTTACAGTACGACGTGCTGATGACCGCTGAGAACCGATGACGAACTCG 300
QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyValAspGluGln 120
Db 301 CCGTGGCCGACCGCCCTCATGAGCCCTGCTGCGGCTTCGTGAAAGGCAAGACGACCAA 360
QY 121 SerArgLeuLeuArgArgThrLeuIleArgTrpAlaAsnLeuGlyAsnValLeuIleLeu 140
Db 361 GCGCGGCTGCTGCGGCGGACGCTCATCCGCTTACGACCTGGGCAACGTGCTCATCTG 420
QY 141 ArgSerValSerThrAlaValTrpValArgPheProSerAlaGlnIleLeuValGlnAla 160
Db 421 CGCAGCGTCAGACCGGACGTCTCAAGCGCTTCCCGACGCCGACGACCTGCTGCAAGCA 480
QY 161 GlyPheMetThrProAlaGluIleValGlyGlnLeuGluValSerLeuProIleAsnMet 180
Db 481 GGCCTTATATGACTCCGGGAGAACCAAGCAATTTGAGAACTGAGCCCTACACCAACATG 540
QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyGlyArg 200
Db 541 TTCTGGGTGCTTACGAGTGTGCTTTCGCAACTGTCAATGAAGCGTGGCTTGAAGGTGCA 600
QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgThrGln 220
Db 601 ATCCGGGACCCCTAATCCGCTCCGACGAGCTGCTGAACGATGAACACTTGGCTACG 660
QY 221 CysGlyIleIleLeuTrpAlaTrpAspTrpIleSerIleProLeuValTrpThrGlnVal 240
Db 661 TGTGGACACTGTATGCTTACGCTGATGATTAATCCACTGTGTATATACAGGTGGTG 720
QY 241 ThrValAlaValTrpSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
Db 721 ACTGTGGCGGTATCAGCTTCTTCTGACTGTGTCTAGTTGGCGGCAAGTTCTGAAACCA 780
QY 261 AlaLeuAlaTrpProGlyIleIleGluLeuAspLeuValAlaProValPheThrPheLeuGln 280
Db 781 GCCAAGGCTTACCTGCGCATGAGCTGAGACTGCTGTGTGCGCTTCAAGTTCCTGCA 840
QY 281 PhePhePheTrpAlaGlyTrpLeuValAlaGluGlnLeuIleAsnProPheGlyGln 300
Db 841 TTTCTTCTTATTTGGCTGGAGGTGGAGGAGCAAGCTCATCAACCCCTTTGAGAG 900
QY 301 AspAspAspPheGluThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeu 320
Db 901 GATGATGATGATTTTATGACCAACTGATTTGTGACAGAAATTTGAGAGGTCTCCCTG 960
QY 321 AlaValAspGluMetIleGlnAspLeuProArgMetGluProAspMetTrpTrpAsnLys 340
Db 961 GCTGTGGATGATGATGACCAAGCAAGCTGCTCGGATGAGAGCGGACATGTACTGAAATTA 1020
QY 341 ProGluProGlnProProTrpTrpAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
Db 1021 CCGAGGACACAGCCCGCTTACAGAGCTGCTTCCGACCAAGTTCGCTGAGGCTCTTAT 1080
QY 361 GluSerThrPheAsnIleSerLeuAsnLysGluGluMetGluPheGlnProAsnGlnGlu 380
Db 1081 GAGTTCACCTTCAACATCAGCTGAAACAAAGAGAGATGAGTTCCAGCCCAATCAGAG 1140
QY 381 AspGluGluAspAlaIleAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerTrpAsp 400

Db 1141 GACGAGGAGTGTCTCAGCGTGCATCATTTGGCCGCTTCTTAGCGCTGCAGTCCCATGAT 1200
Qy 401 HishisProxofArgAlaAsnSerArgThrLysLeuLeuTTPProLysArgLysSerLeu 420
Db 1201 CACATCTCTCCAGGGCAAACTCAAGACCAAACTACTGTGCCCCAAGAGGAAATCCCTT 1260
Qy 421 LeuHiseGluGlyLeuProLysAsnHisLysValAlaLysGlnAsnValLysGlyGlnGlu 440
Db 1261 CTCACGAGGGGCGTCCCAAAAACACAAAGGACGACAAAGAAAGCTTAAGGGCCAGGAA 1320
Qy 441 ArgAsnLysValATTPPLysLeuLysValAlaLysAlaPheLysSerGlyProLeuTyrGln 460
Db 1321 GACAAACAAGGCGCTGGAAGCTTAAGCTGTGACCGCTTCAAGTCTGCCCTCACTGATCAG 1380
Qy 461 ArgProGlyTyrTyrSerAlaProGlnThrProLysSerProThrPhePhePro 480
Db 1381 AGGCGAGGCTACTACAGTGCACACAGAGCGCCCTCAGCCCACTCCCATGTTCTTCCCC 1440
Qy 481 LeuGluProSerAlaProSerLysLeuHisSerValThrGlyLeuAspThrLysAspLys 500
Db 1441 CTAAACAACATCAGCGCCCTCAAGCTTCAAGTGTCAAGGCAATGACACAAAGACAA 1500
Qy 501 SerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLysSerGluSerAsp 520
Db 1501 AGCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGGCGAT 1560
Qy 521 GlyValLeuMetGluHisProGluValSerGlnValArgValLysThrValGluPheAsn 540
Db 1561 GGGGCGCTTGTAGGACACCCAGAAATATCTCAAGTGAAGAGAGAAACGTGTGAGTTTAC 1620
Qy 541 LeuThrAspMetProGluLeuLeuProGluAsnHisLysLeuLysGluProLysGluGlnSerPro 560
Db 1621 CTGACGGATATGCCAGATCCCGAAATATCACTCAAAAGAACTTTGAAACATTCACCA 1680
Qy 561 ThrAsnLeuHisThrThrLeuLysAspHisMetAspProTyrTTPAlaLeuGluAsnArg 580
Db 1681 ACCAACTACACACTACACTCAAGATCATCATGATCTTATTTGGGCGCTTGAAAACAGG 1740
Qy 581 AspGluAlaHisSer 585
Db 1741 GATGAGCAGCATCC 1755

RESULT 4
AF073501 2171 bp mRNA linear PRI 02-SEP-1998
LOCUS Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA,
DEFINITION complete cds.
ACCESSION AF073501 GI:3511241
VERSION AF073501.1
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2171)
Stohr, H., Marguardt, A., Rivera, A., Cooper, P.R., Nowak, N.J.,
Shows, T.B., Gerhard, D.S. and Weber, B.H.,
A gene map of the Best's vitelliform macular dystrophy region in
chromosome 11q12-q13.1
JOURNAL Genome Res. 8 (1), 48-56 (1998)
MEDLINE 9445487
PUBMED 9445487
REFERENCE 2 (bases 1 to 2171)
Marguardt, A., Stohr, H., Passmore, L., Kraemer, F., Rivera, A. and
Weber, B.H., P.
Direct SubMISSION
JOURNAL Submitted (22-JUN-1998) Human Genetics, Universitt, Biozentrum, Am
Hubland, Wuerzburg 97074, Germany
FEATURES
SOURCE 1..2171
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

gene /chromosome="11"
/map="11q12-q13.1"
/tissue type="retinal pigment epithelium"
1..2171
/gene="VMD2"
89..1846
/gene="VMD2"
/note="Best disease"
/product="vitelliform macular dystrophy protein"
/protein_id="AAC33766.1"
/db_xref="GI:3511242"
/translation="MTITTSQVANAELGFSRLLLCWRSIYKLLYGEFLIFLLCY
IIRFIRLALTEEQLEMEKLTLYCDSYIOLIPISFVLGFVTLVVRMNQYENLP
PDRLSLVSGFVEGKDEQRLRRITLIRYANLGVLLRSVSTAVRPSAOHLVOA
GPMPLAHKOLEKLSLPHNMFVPMVPMANLSMAWLGRIIDPILLOSILNEMNTLR
TQCHLYAYHMISIPLYTTOVTVAYVSFFLTCLVGGQPLNPAKAYGHELDVVPF
TFLOFFPYVMWLKVAEQLINPGEEDDFFTNMIVDSNLQVSLAIVDEMHODLPMBP
DMWNKPEQPPRYTAASAOFRRASFMGSTENISLKNKEMEPQNPQDEEDNAHGIOR
FLGIQSDHHPPRANRSRTKLMLPKESELHSEGLPKNKAQKQNVGQEDNAWMLKAY
DAFSAFLYORPGYSAPOPLASPTPMPFLEPSPAPSLKHSVTGIDTRKSLKTVSSG
AKSFELSSDGALEMBEPVSQVRRTVEFNLTDMPEIENHLKEPLDEGSPTHHT
LKHMDPYMALENDENAS"

ORIGIN

Alignment Scores:

Pred. No.:	2,22e-266	Length:	2171
Score:	3110.00	Matches:	583
Percent Similarity:	99.66%	Conservative:	0
Best Local Similarity:	99.66%	Mismatch:	2
Query Match:	99.68%	Indels:	0
DB:	9	Gaps:	0

US-09-622-964a-3 (1-585) x AF073501 (1-2171)

Qy 1 MetThrIleThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db 89 ATGACATCACTTACACAAAGCAAGTGGCTAATCCCGCTTAGGCTCCTTCCCGCCTG 148
Qy 21 LeuLeuGlySerPheArgGlySerIleTyrLysLeuLeuTyrGlyGluPheLeuLysPheLeu 40
Db 149 CTGCTGGCTGGCGGGGACAGCATCTACAGCTGCTAATAGCGAGTCTTAATCTTCCG 208
Qy 41 LeuCySerTyrTyrIleIleArgPheIleTyrArgLeuAlaLeuThrGluGlnGlnLeu 60
Db 209 CTCTGCTACTACATCAACCGCTTATTTATGCTGGCCCTCAAGGAAGAACACAGCTG 268
Qy 61 MetPheGluLysLeuThrLeuTyrCyAspSerTyrIleGlnLeuIleProLysSerPhe 80
Db 269 ATGTTGAAAGACGACGCTGTATTGGGACAGCTACATCCAGCTCATCCCATTTCCCTTC 328
Qy 81 ValLeuGlyPheTyrValThrLeuValAlaThrArgTTPTPanGlnTyrGluAsnLeu 100
Db 329 GTGCTGGGCTTCTTACGAGACCTGCTGTGACCCGCTGTGGGAACAGTACGAAACCTG 388
Qy 101 ProThrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGln 120
Db 389 CCGTGGCCGACCCCTCATGAGCTGTGTGGGCTTCGTCAAGGCAAGACGAGCA 448
Qy 121 SerArgLeuLeuArgArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeuLeu 140
Db 449 GCGCGGCGTGGCGGGCGGCGCATCCGCTACGCCAAACCTGGGCAACGTGCTCATCTCG 508
Qy 141 ArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAla 160
Db 509 GCGACGCTCAGCAGCGGACGCTTACAGCGCTTCCCGCGCCGACCTGTGTCCAAACA 568
Qy 161 GlyPheMetThrProAlaGluHisLysGlnLeuGluLysLeuSerLeuProHisAsnMet 180
Db 569 GGCCTTATGACTCCGGCAACACAGAGCTTGAAGAACTGACCTTACCAACAACATG 628
Qy 181 PheThrValProTTPValTTPPheAlaAsnLeuSerMetLysValATTPLeuGlyGlyArg 200

Db 629 TTCTGGGTGCTGGGTGGTGGTTTGCACACTGTCATGTAAGCGCTGGCTTGAGAGTCA 688
 Qy 201 ILeArgAspProIleLeuLeuGlnSerLeuLeuEngIUmEAsnThrLeuArgThrGln 220
 Db 689 ATCCGGAGCCCTATCTCTCTCCAGAGCCTGCGAAGAGTAAGAACCTTGCGATCACTAG 748
 Qy 221 CysGlyHisLeuYrAlaYrAspTrpIleSerIleProLeuValYrThrGlnValVal 240
 Db 749 TGTGGACACCTGTATGCTACGACTGATTTAGTATCCACAGTGTCTATACACAGTGGTGTG 808
 Qy 241 ThrValAlaValYrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
 Db 809 ACTGGGGGGGTGATACGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 868
 Qy 261 AlaValAlaYrProGlyHisGlnLeuAsnLeuValProValPheThrPheLeuGln 280
 Db 869 GCCAAGGCTTACCTGGCCATGAGCTGAGCCTCGTGTGCGCGCTTACAGTCTCTGAG 928
 Qy 281 PhePhePheYrValGlyTrpLeuYsValAlaGlnGlnLeuLeuAsnProPheGlyGln 300
 Db 929 TTCTTCTTCTATGTTGGCTGGCTGAGGTGGCAGAGCAGCTCATCAACCCCTTGAGAG 988
 Qy 301 AspAspAspAspPheGlnThrAsnTrpIleValAspArgAsnLeuGlnValSerLeu 320
 Db 989 GATGATGATGATTTTGAAGACCAACTGATGTCGACAGAAATTTGCAAGGTGCTCCCTGTTG 1048
 Qy 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetYrTrpAsnLys 340
 Db 1049 GCTGTGATGATGATGACACAGACCTGCTGAGTGGAGCGGACAGTACTCGAATAAG 1108
 Qy 341 ProGluProGlnProProYrThrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
 Db 1109 CCCGAGCAGCAGCCCTTACACAGCTGCTCCGCCAGCTCCGTCAGAGCTCCTTTATG 1168
 Qy 361 GlySerThrPheAsnIleSerLeuAsnLysGlnGlnMetGluPheGlnProAsnGlnGln 380
 Db 1169 GGCCTCCACTTCAATCAGCTGAAACAAAGAGAGATGAGTTCACGCCCAATCAGAG 1228
 Qy 381 AspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
 Db 1229 GACGAGAGATGCTCAGCGTGGCATCTTGGCGCTTCTTGAAGCTGCACTCCCATGAT 1288
 Qy 401 HisHisProProArgAlaAsnSerArgThrLysLeuLeuTrpProLysArgLysSerLeu 420
 Db 1289 CACCATCCTCCCAAGGCAAACTCAAGGACCAAACTAGTGTGCCCAAGAGGAATCCCTT 1348
 Qy 421 LeuHisGlnGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGln 440
 Db 1349 CTCACAGAGGCGCTGCCCAAAACCAAGGACCAAAAGAACTTGAAGGCGCAGAA 1408
 Qy 441 AspAsnLysAlaTrpLysLeuLysAlaValAspAlaPheLysSerGlyProLeuTrpGln 460
 Db 1409 GACAAACAGGCTGGAAGCTTAAGGCTGTGAGCCCTTCAAGTGTGCCCACTGTATCAG 1468
 Qy 461 ArgProGlyYrYrSerAlaProGlnThrProLeuSerProThrPrometPhePhePro 480
 Db 1469 AGGCGAGGCTACTACAGTGCACCAAGACACCCCTCAGGCCCACTCCCATGTTCTTCCC 1528
 Qy 481 LeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLys 500
 Db 1529 CTAAACACATCAGGCGCTCAAAAGCTTCAAGTGTCAAGGCAATGACACCAAAAGACAAA 1588
 Qy 501 SerLeuYrThrValSerSerGlyAlaLysLysSerPheGluLeuLeuSerGluSerAsp 520
 Db 1589 AGCTTAAGACTGTGAAGTTCGCGGCCCAAGAAAGTTTGAATTCCTCTCAGAGAGCGAT 1648
 Qy 521 GlyAlaLeuMetGluHisProGluValSerGlnValArgArgLysThrValGluPheAsn 540
 Db 1649 GGGGCGCTTGATGAGCACCAAGAGTATCTCAAGTGAAGAGAAACCTGAGAGATTAAAC 1708
 Qy 541 LeuThrAspMetProGluLysProGluAsnHisLeuLysGlnProLeuGluGlnSerPro 560
 Db 1709 CTGACGGATATGCGAGAGATCCCGAAATCACTCAAAAGAACTTTGGAACAATCACCA 1768

Qy 561 ThrAsnIleHisThrThrLeuLysAspHisMetAspProYrTrpAlaLeuGluAsnArg 580
 Db 1769 ACCAACAATACACTACATCAAGATCACTGATCTTATTTGGGCTTGAAACAGG 1828
 Qy 581 AspGluAlaHisSer 585
 Db 1829 GATGAAGCACATTCC 1843

RESULT 5

AF057169 2210 bp mRNA linear PRI 17-OCT-1998
 Homo sapiens beastrophin (VMD2) mRNA, alternatively spliced product,
 complete cds.

AF057169
 AF057169.1 GI:3335158

ACCESSION

AF057169.1 GI:3335158

VERSION

AF057169.1 GI:3335158

KEYWORDS

complete cds.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 2210)

AUTHORS

Petrukhin, K., Koishi, M.J., Bakall, B., Li, W., Xie, G., Markwell, T.,
 Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M.,
 Bergen, A.A., McGarvey, Dugan, V., Figueroa, D., Austin, C.P.,
 Metzger, M.L., Caskey, C.T. and Madefius, C.
 Identification of the gene responsible for Best macular dystrophy
 Nat. Genet. 19 (3), 241-247 (1998)

TITLE

Submitted (03-APR-1998) Human Genetics, Merck Research
 Laboratories, West Point, PA 19486, USA

JOURNAL

Location/Qualifiers

FEATURES

1..2210

source

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 1..2210
 /gene="VMD2"
 105..1862
 /gene="VMD2"
 /note="Best's macular dystrophy gene; BMD, alternatively
 splice isoform"
 /product="beastrophin"
 /protein_id="AAC64343.1"
 /db_xref="GI:3335159"
 /translation="MTTITTSQVANKRLSFSRLICWRSYIKLYGEPFLFLCY
 IIRFIRLALTEBOQLMEFKTLTCDYIQLPISVLEFVYTLVVTWMDQENLPW
 PDLMLVSGFVGEKDEOGLRLRLIRVANIENVILISVSAVYKRPSPQHLVOA
 GEFTPAHEKOLEKLSLPHNFMVPMWPFANLSMKAVLIGRIDPILLOSLINEMTLR
 TOCGHYAYDWISIPLYOVTVAVSFELICVGRPLNPAKAYGHELDLVVVF
 TRLQPFVYQYKVAEOLINPGEDDDDEFTWIVDRNLQVSLIADNEMHODLPREP
 DMWPKPEPQPTTASAOFRASFWGTFNLSLKNEMEFOQNDDEDAHAGITGR
 FLGLOSHDRHPANSRTYLLMPKRESLHGLPNXKAKAQNVRQEDNKKMKLKA
 DAFKSLPDRPGYSAQPTPLSPFMFPLPSASVSKLSYTGIDTKRKSITVSSG
 AKKSFLLSEDOALMEHPEVGSQVRKTVFNLTPMPEIPENHLKEPLBQSPNHTT
 LNDHNDPYALNREDAHS"

gene

1..2210

CDS

105..1862

ORIGIN

Alignment Scores:

Pred. No.: 2,28e-266 Length: 2210
 Score: 3110.00 Matches: 583
 Best Similarity: 99.66% Conservative: 0
 Best Local Similarity: 99.66% Mismatches: 2
 Query Match: 99.68% Indels: 0
 DB: 9 Gaps: 0

US-09-622-964a-3 (1-585) x AF057169 (1-2210)

Qy 1 MetThrIleThrTyThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
 Db 105 ATGACCATCACTTACACAGCCAGTGGCTAAATGCCCTTAAGCTCTTCTCCGCGCTG 164
 Qy 21 LeuLeuCyTrpArgGlySerIleTyrlsLeuLeuTyrglyGlnPheLeuIlePheLeu 40
 Db 165 CTGTGTGTGTGGCGGGGAGCATCTACAAAGCTGCTAATGTGGAGATTCTTAATCTTCCG 224
 Qy 41 LeuCyTrpTyrlleIleArgPheIleTyrlsArgLeuAlaLeuThrGlnGlnGlnLeu 60
 Db 225 CTCTGCATCACTACATCCGCTTAAATTTATGTGGCTGGCCCTCAGAGAAACAAACAGTGTG 284
 Qy 61 MetPheGlnLysLeuThrLeuTyrcysAspSerTyrlleGlnLeuIleProIleSerPhe 80
 Db 285 ATGTTTGAGAACTGACTCTGTATTTGGACAGCTACATCCAGCTCATTCCTTCTTCC 344
 Qy 81 ValLeuGlyPheTyrlsValThrLeuValAlaThrArgTrpTrpAsnGlnTyrgluAsnLeu 100
 Db 345 GTGCTGGGCTTCTACGTGAGCGCTGGTGTGACCCGCTGGTGGAAACAGTACAGAACCTGTG 404
 Qy 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGlnGlyLysAspGlnGln 120
 Db 405 CCGTGGCCCGACCCGCTCATGAGCTGGTGTGTGGGCTTCTGTGAAAGGCAAGAGCAAGCAA 464
 Qy 121 SerArgLeuLeuArgArgThrLeuIleArgTrpAlaAsnLeuGlyAsnValLeuIleLeu 140
 Db 465 GGGCGGCTGGCGGCGGCGACGCTCATCGCTACGCAACCTGGGCGCAAGAGCTCATCTGTG 524
 Qy 141 ArgSerValSerThrAlaValAlaTyrlsLeuArgPheProSerAlaGlnIleLeuValGlnAla 160
 Db 525 CGCAGCCTCAGCAGCAGCTTACAGAGGCTTCCCGAGCGCCAGCAGCAGCTGTGCAAGCA 584
 Qy 161 GlyPheMetThrProAlaGlnIleGlyGlnLeuGlyLysLeuSerLeuProIleAsnMet 180
 Db 585 GGGCTTTATGACTCCGGCAGAAACAAAGCAGTGGAGAAACTGAGCTTACACACAACTAG 644
 Qy 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLysGlyArg 200
 Db 645 TTCTGGGCGCTGGTGGTGGTGGTGGCAACTGTCAAGAGGGGTGGCTGGAGGCTGCA 704
 Qy 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGln 220
 Db 705 ATCCGGAACCTTATCCGCTCCAGAGCTGCTGAACAGATGAACACTTGCATATCTCAG 764
 Qy 221 CysGlyIleLeuTyrlsAlaTyrcysTrpIleSerIleProLeuValTyrlsGlnValAla 240
 Db 765 TGTGGAACACTGTATGCTTACGACTGATTAATCCACTGGTGTATACACAGGTGGTG 824
 Qy 241 ThrValAlaValTyrcysPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
 Db 825 ACTGTGGCGGTGTACAGCTTCTCTGTGCTGTGCTGTGAGGCGGAGTCTCTCAACCCA 884
 Qy 261 AlaIysAlaTyrcysGlyIleGlnLeuAspLeuValAlaProValPheThrPheLeuGln 280
 Db 885 GCCAAGGCTTACCTGGCCATGAGCTGACCTCGTTGTGCCCTTCTTACCTTCTGAG 944
 Qy 281 PhePhePheTyrlsValGlyTrpLeuLysValAlaGlnGlnLeuIleAsnProPheGlyGln 300
 Db 945 TTCTTCTTCTATGTTGGCTGTGAGGTGGCAGAGCGCTCATCAACCCCTTTGGAGAG 1004
 Qy 301 AspAspAspAspPheGlyThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeu 320
 Db 1005 GATCATGATGATTTTGGACCAACTGATGTGTGCAAGAGATTTTGCAGGTCCTCCCTGTG 1064
 Qy 321 AlaValAspGlnMetHisGlnAspLeuProArgMetGlnProAspMetTyrlsTrpAsnLys 340
 Db 1065 GCTGTGTGTGAGTGCACCAAGACCTGCTGTGATGTGAGCGGCAATGTACTGTGAATAG 1124
 Qy 341 ProGlnProGlnProProTyrlsThrAlaIleSerAlaGlnPheArgArgAlaSerPheMet 360
 Db 1125 CCGAGGCCACAGCCCTTACACAGCTGCTCCGCGCAAGTTCGCTGAGAGCTTCCCTTATG 1184

Qy 361 GlySerThrPheAsnIleSerLeuAsnLysGlnLysMetGlnPheGlnProAsnGlnGln 380
 Db 1185 GGTCCACCTTCAACATCACCTGAAACAAGAGAGATGAGTTGCCAGCCCAATCAGAG 1244
 Qy 381 AspGlnGlnAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
 Db 1245 GACAGAGAGATGTCTACGCTGGCATTTGGCGGCTTCTTCAAGGCTGCAGTCTCATGAT 1304
 Qy 401 HisIleProProArgAlaAsnSerArgThrTyrlsLeuLeuTrpProLysArgLysLeu 420
 Db 1305 CACCATCTCCACAGGGCAAACTCAAGSACAACTTGTGGCCCAAGAGGAAATCCCTT 1364
 Qy 421 LeuHisGlnGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGln 440
 Db 1365 CTCACAGAGGGCCCGCCCAAAACCAACAGGCGCAACAGACGTTAGGGGCGCAGAA 1424
 Qy 441 AspAsnLysAlaTrpLysLeuLysValAlaAspAlaPheLysSerGlyProLeuTyrlsGln 460
 Db 1425 GACAAACAGGCTGGAAGCTTAAAGCTGTGACGCTTCAAGTCTGCCCACTGTATCAG 1484
 Qy 461 ArgProGlyTyrlsTyrcysAlaProGlnThrProLeuSerProThrProMetPhePhePro 480
 Db 1485 AGGCGAGGCTTACATCACTGCCCCACAGAGCCCTCAGCCCACTCCCATGTCTTCC 1544
 Qy 481 LeuGlnProSerAlaProSerLysLeuHisSerValThrGlyIleAspThrLysAspLys 500
 Db 1545 CTAGAACCATCAGGCGCGCTCAAGCTTCAACAGTCTCAGAGCATAGACACCAAGACAA 1604
 Qy 501 SerLeuLysThrValSerSerGlyAlaLysLysSerPheGlnLeuLeuSerGlnSerAsp 520
 Db 1605 AGCTTAAAGACTGAGTGTCTGGGGCCAGAGAAAGTTTGAATGTCTCAGAGAGCGAT 1664
 Qy 521 GlyAlaLeuMetGlnHisProGlnValSerGlnValAlaArgArgLysThrValGlnPheAsn 540
 Db 1665 GGGGCTTGTGAGGACCCAGAGATCTCAAGTGTGAGAGGAAACTGTGAGATTTAAC 1724
 Qy 541 LeuThrAspMetProGlnIleProGlnAsnHisLeuLysGlnProLeuGlnGlnSerPro 560
 Db 1725 CTGACGATTTGCGAGATCTCCGAAATCATCTCAAAAGAACTTTGGAACATTCACA 1784
 Qy 561 ThrAsnIleHisThrThrLeuLysAspHisMetAspProTyrlsAlaLeuGlnAsnArg 580
 Db 1785 ACCAATATCACACTACACTCAAGATCAGATGATCTTATTTGGCTTGGAAACAGG 1844
 Qy 581 AspGlnAlaHisSer 585
 Db 1845 GATGAAGCACATTC 1859

RESULT 6
 LOCUS CQ729633 2170 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 15567 from Patent WO2068579.
 ACCESSION CQ729633
 VERSION CQ729633.1 GI:42301252
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 Kites, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 Patent: WO 02068579-A 15567 06-SEP-2002;
 JOURNAL PE Corporation (NY) (US)
 FEATRES Location/Qualifiers
 SOURCE 1..2170
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	5,66-264	Length:	2170
Score:	3083.00	Matches:	582
Percent Similarity:	99.49%	Conservative:	0
Best Local Similarity:	99.49%	Mismatches:	3
Query Match:	98.81%	Indels:	1
DB:	6	Gaps:	0

US-09-622-964a-3 (1-585) x CQ729633 (1-2170)

QY 1 MetThrIleThrYrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
DB ATGACCACTCACTTACACAGCCAAAG-GCTAATGCCCGCTTAGGCTCTTCTCCGCTG 147
QY 21 LeuLeuCystrPArgGlySerIleTyrlYsLeuLeuTyrglyGluPheLeuIlePheLeu 40
DB CTGCTGAGCTGGCGGGGCGCATCTTCAAGCTGCTAATAGCGAGTCTTAACTTCTCTG 207
QY 41 LeuCystrYrYrIleIleArgPheIleTyrlArgLeuAlaLeuThrGluGlnGlnLeu 60
DB CTCTGCTACTACATCATCCGCTTATTATAGCTGGCCCTCAGGAAGAACACAGCTG 267
QY 61 MetPheGluYsLeuThrLeuTyrlCysAspSerTyrlIleGlnLeuIleProIleSerPhe 80
DB ATGTTTGAAGAACTGACTCTGTATTGGCAGACATCCAGCTCATCCCATTTCTCTTC 327
QY 81 ValLeuGlyPheTyrlValIleThrLeuValValThrArgTrpTrpAsnGlnTyrglyAsnLeu 100
DB GTGCTGGAGCTTTTACGAGCCCTGTGTGTGACCCGCTGGTGGAAACAGTACAGAACTG 387
QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyYsAspGluGln 120
DB CCGTGGCCCGACCGCTCATGAGCCGTGTGTGGCTTCGTGGAAGCAAGACAGACAA 447
QY 121 SerArgLeuLeuAlaGlnThrLeuIleArgTyrlAlaAsnLeuGlyAsnValLeuIleLeu 140
DB GGCGGCTGCTGCGCGCAGCCTCATCCGCTACCCAGCCAACTGGGCAACGTGCTATCTG 507
QY 141 ArgSerValSerThrAlaValTyrlYsArgPheProSerAlaGlnIleSLeuValGlnAla 160
DB CGACGCTACAGACCGGAGTTCACAGGCGCTTCCCGAGCCCGCCAGCACTGGTTCAGACA 567
QY 161 GlyPheMetThrProAlaGlnIleSLeuGlnLeuGlyYsLeuSerLeuProIleAsnMet 180
DB GGCTTTATGACTCCGGCAGAACACAGAGATTGGAAGAACTGACCTTACACACACATG 627
QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetYsAlaTrpLeuGlyYsArg 200
DB TTCTGGGTGCCCTGGGTGTGTGTTCACAACTGTCAATGAAGGTGGCTTGGAGGTGCA 687
QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGln 220
DB ATCCGGAACCTTATCTGCTCTCAAGGCTGCTGAACAGATGAACACCTTGCATCTCAG 747
QY 221 CysGlyHisLeuTyrlAlaTyrlAspTrpIleSerIleProLeuValTyrlThrGlnValVal 240
DB TTGTGACACCTGTATGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 807
QY 241 ThrValAlaValTyrlSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
DB ACTGTGGCGGTGTACACTTCTCTCTGCTAGCTTGTGGGCGGAGTTTCTGAAACCA 867
QY 261 AlaValAlaTyrlProGlyHisGlnLeuAspLeuValAlaProValPheThrPheLeuGln 280
DB GCCAAGGCTTACCTGGCAGTACAGCTGACTCTGTGGCCGCTTCTTCACTTCTCAG 927
QY 281 PhePhePheTyrlValGlyTrpLeuYsValAlaGluGlnLeuIleAsnProPheGlyGlu 300
DB TTCTTCTTCTATGTGTGGCTGGAAGTGGCAGAGCCTCAACACCTTTTGGAGAG 987
QY 301 AspAspAspAspPheGluThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeu 320

DB 988 GATGATGATGATTTTGAACCAACTGGATTGTGACAGGAAATTTGACAGGTCTCCCTGTTG 1047
QY 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrlTrpAsnYs 340
DB GCTGTGATGATGATGACACGAGACCTGCTCGATGAGCGGACATGATCTGGAATAG 1107
QY 341 ProGluProGlnProProTyrlThrAlaAlaSerAlaGlnPheArgAlaSerPheMet 360
DB CCCGAGCCAGGCCCTTCAACAGCTGTCTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTT 1167
QY 361 GlySerThrPheAsnIleSerLeuAsnYsGluGlnMetGluPheGlnProAsnGlnGlu 380
DB GGCTTCACTTCAACATCACTGATGAAACAGAGAGATGAGATTCAGCCCAATCAGAG 1227
QY 381 AspGluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
DB GACGAGAGAGATGCTACGCTGGCATCATTTGGCGCTTCTTACGCTGAGCTCCATGAT 1287
QY 401 HisHisProProArgAlaAsnSerArgThrYsLeuLeuTrpProYsArgGluSerLeu 420
DB CACCATCTCCAGGCGCAACTCAGAGCCAACTACTGTGGCCCAAGAGGAAATCCCTT 1347
QY 421 LeuHisGluGlyLeuProYsAsnHisSLeuAlaAlaGlnAsnValArgGlyGlnGlu 440
DB CTCACAGAGGCTGCTCCCAAAACACAGAGCAGCCAAACAGACGTTAGAGGCCAGAAA 1407
QY 441 AspAsnYsAlaTrpYsLeuYsAlaValAspAlaPheYsSerGlyProLeuTyrlGln 460
DB GACACAGAGGCTCGAAGCTTAAAGCTGTGAGCGCTTCAAGTCCGCTCAGCTATCAG 1467
QY 461 ArgProGlyTyrlYsSerAlaProGlnThrProLeuSerProThrProMetPhePhePro 480
DB AGGCGAGGCTTACAGTGTGCCCAAGCCGCTCAGCCCTCAGCCCTCAGCTTCTTCCC 1527
QY 481 LeuGlnProSerAlaProSerLeuHisSerValThrGlyIleAspThrYsAspYs 500
DB CTAGAACATCAGAGCGGCTTAAAGCTTCAACAGTGTCAAGAGCATAGACCAAGACAA 1587
QY 501 SerLeuYsThrValSerSerGlyAlaYsYsSerPheGluLeuLeuSerGluSerAsp 520
DB AGCTTAAAGCTGAGTGTGGGGCCAAAGAAATTTGAAATTTGCTCAGAGAGGAT 1647
QY 521 GlyAlaLeuMetGluHisProGluValSerGlnValArgArgYsThrValGluPheAsn 540
DB GGCGCTTGTAGAGCACCCAGAAATATCTCAAGTGGAGGAAATCTGGAGTTTAC 1707
QY 541 LeuThrAspMetProGluIleProGluAsnHisSLeuYsGluProLeuGluGlnSerPro 560
DB CTAGCGATATGCGAGAGATCCCGAAATCACTTCAAGAACCTTTGAAACATCAACA 1767
QY 561 ThrAsnIleHisThrThrLeuYsAspHisMetAspProTyrlTrpAlaLeuGluAsnArg 580
DB ACCAACAATCAACATCACTCAAAAGATCAATGATCTTATTTGGGCTTGGAAAACAGG 1827
QY 581 AspGluAlaHisSer 585
DB 1828 GATGAAGCACTTCC 1842

RESULT 7
BD136721 2429 bp DNA linear PAT 18-SEP-2002
LOCUS BD136721
DEFINITION Best's macular dystrophy gene.
ACCESSION BD136721
VERSION BD136721.1 GI:23231666
KEYWORDS JP 2002504559-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2429)
AUTHORS Petrukhin,K., Caskey,T.C., Metzker,M. and Wadelius,C.
TITLE Best's macular dystrophy gene
JOURNAL Patent: JP 2002504559-A 3 12-FEB-2002;

COMMENT MERCK & CO INC, CLAES WADELIUS
OS Homo sapiens (human)
PN JP 2002504559-A/3
PD 12-FEB-2002
PF 22-FEB-1999 JP 2000533447
PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUHIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI
WADELIUS
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P13/34, C12Q1/68// PC
C12P21/08,
PC C12N5/00, C12N15/00
CC Best's macular dystrophy gene
FH Location/Qualifiers
FT Key 1..2429
FT source /organism='Homo sapiens (human)'.
FEATURES
source 1..2429
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Alignment Scores:
Pred. No.: 9,32e-261 Length: 2429
Score: 3047.50 Matches: 583
Percent Similarity: 89.28% Conservative: 0
Best Local Similarity: 89.28% Mismatches: 2
Query Match: 97.68% Indels: 69
DB: Gaps: 1
US-09-622-964A-3 (1-585) x BD136721 (1-2429)
QY 1 MetThrIleThrTrpThSerGlnValAlaAspAlaArgLeuGlySerPheSerArgLeu 20
DB 105 ATGACCACTACTTACACAAAGCAAGTGGCTAATGCGCTTAAGCTCTTCTCCGCGCTG 164
QY 21 LeuLeuCybTPAAGlySerIleTyxLeuLeuTyxGlyPheLeuIlePheLeu 40
DB 165 CTGGCTGCTGGCGGGGAGCATCTACAAAGCTGCAATGGCGAGTTTAAATCTTCTCG 224
QY 41 LeuCybTyxIleIleIleArgPheIleTyxArgLeuAlaLeuThrGlyGlnGlnLeu 60
DB 225 CTGTGCTACTACATCATCCGCTTATTTATAGCTGGCGCTTACGGAAGAACACAGCTG 284
QY 61 MetPheGlyValLeuThyLeuTyxGlyAspSerTyxIleGlnLeuIleProIleSerPhe 80
DB 285 ATGTTTGAGAACTGACTCTGTATTGCGACAGCTACATCCAGCTCATCTCCATTCTTC 344
QY 81 ValLeuGlyPheTyxValIleThyValValIleThrArgTrpTPAAspGlnTyxGlyAsnLeu 100
DB 345 GTGCTGGGCTTCTAGTACCGCTGTGACCGGCTGTGGAAACAGTACAGAAACCTG 404
QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGlyGlyAspGlyGln 120
DB 405 CCGTGGCCCGACCGGCTCATGAGCTGGTGTGGGCTTCGTGAAGGCAAGACGAGCAA 464
QY 121 SerArgLeuLeuAlaGlyGlyThyLeuIleArgTyxAlaAsnLeuGlyAsnValLeuIleLeu 140
DB 465 GCGCGGCTGTCGCGCGACGCTCATTCGCTACGCAACCTGGGCAACCTTCTCATCTCG 524
QY 141 ArgSerValSerThrAlaValIleTyxValArgPheProSerAlaGlnIleValValGlnAla 160
DB 525 CGCAGCGTCAGACCGAGTCTTACAGGCGCTTCCCGACGCCACAGCACTTGTGCAAGCA 584
QY 161 GlyPheMetThrProAlaGlyIleGlyGlnLeuGlyValLeuSerLeuProIleAsnMet 180
DB 585 GGCCTTTATGACTCCGGCAGAACACAAAGATTGGAGAACTGAGGCTTACACACAACATG 644
QY 181 PheThrValProTrpValIleThyPheAlaAsnLeuSerMetValIleTrpLeuGlyValArg 200
DB 645 TTTCGGGCGCCCTGGGTGTGGTTTGGCAACCTGTCAATGAAGGCGTGGCTTGGAGGTGCA 704
QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGln 220

DB 705 ATCCGGAGCCCTATCTGCTCCAGAGCCTGCTGAACAGATGAACACCTTGCGATCTAG 764
QY 221 CyeGlyHisLeuTyxAlaTyxAspTrpIleSerIleProLeuValIleThyGlnValVal 240
DB 765 TGTGAGACCTGTAATGGCTTACAGACTGAGTATGATCCACATGGTGTATACAGAGTGTG 824
QY 241 ThrValAlaValTyxSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
DB 825 ACTGTGGCGGTACAGCTTCTTCTGACTGTGTAGTGGCGGCACTTCTGAAACCA 884
QY 261 AlaValAlaTyxProGlyHisGlyLeuAspLeuValProValPheThrPheLeuGln 280
DB 885 GCCAAGGCTTACCTGGCGCATAGACTGAACTGCTGTGGCCCGCTTCAACCTTCTCGCAG 944
QY 281 PhePhePheTyxValGlyTrpLeuVal----- 290
DB 945 TTCTTCTTCTAATGTGTGGCTGGCTGAAGGT-GGGCTTCTCAAGGCGCTGCGGCTGAG 1003
QY 290 ----- 290
DB 1004 GCATGGCCAGAGGGGTATGCGCAGCTGCTTGAACAGAGATGCAGTGTCAAGAAAG 1063
QY 290 ----- 290
DB 1064 GAAGTCTCACGGGTAGAAAGCAGCAGCGGTGGGCGACACCTGTATCCAGCTTAC 1123
QY 291 ----- 1123
DB 1124 TCGGAGAGCTGAGGACAGAGAAATGCTTGAACCCGGAAGCGGAGTGTGGTGGCGAGAG 1183
QY 293 GlnLeuIleAsnProPheGlyGlyAspAspAspPheGlyThrAsnTrpIleValAsp 312
DB 1184 CAGCTCATCAACCCCTTTGAGAGAGATGATGATATTTTGAACCACTGGATGTGTGAC 1243
QY 313 ArgAsnLeuGlnValSerLeuLeuValAspGlyMetHisGlnAspLeuProArgMet 332
DB 1244 AGGAATTTGAGAGGTGCTCGTGTGGTGTGGATGAGATGACCAAGACCTTCTCGGATG 1303
QY 333 GlnProAspMetTyxTrpAsnValProGlyProGlyProProTyxThrAlaIleSerAla 352
DB 1304 GAGCGGAGCATGTACTGTAATTAACCCGAGGCAAGCCCTTACAGAGCTTCCGCC 1363
QY 353 GlnPheArgArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnValGlyGln 372
DB 1364 CAGTTCGCTGAGGCTCTCTTATGTGGCTCCACCTTCAACATCAGCCTGAACAAAGAGAG 1423
QY 373 MetGlnPheGlnProAspGlnGlyAspGlyGlyAspAlaIleValAlaGlyIleGlyArg 392
DB 1424 ATGAGATTCCAGCCCAATCAGAGAGAGAGAGAGATGCTACGCTGGCATCATTTGGCGCG 1483
QY 393 PheLeuGlyLeuGlnSerHisAspHisIleProProArgAlaAsnSerArgThyLeuLeu 412
DB 1484 TTCTTAAGCTGAGCTCCCATGATCACCATCTTCCAGGGCAAACTCAAGAGCAACCACTA 1543
QY 413 LeuTrpProIleAspGlySerLeuLeuHisGlyGlyLeuProIleValAsnHisIleValAla 432
DB 1544 CTGTGCGCCCAAGAGAGAAATCCCTTCTCACAGAGGCGCTGGCCAAAACCAAGAGAGCC 1603
QY 433 LysGlnAsnValArgGlyGlnGlyAspAsnValIleTrpValLeuValAlaValAspAla 452
DB 1604 AAACAGAACTGTAGGGGCGAGAGAGACAAAGGCTCGGAAGCTTAAAGGCTGTGAGCGCC 1663
QY 453 PheIleSerGlyProLeuTyxArgProGlyTyxTyxSerAlaProGlnThrProLeu 472
DB 1664 TTCAAGTCTGGCCCACTGTATCAGAGGCGCAGGCTACTACAGTGGCCCAAGAACCCCTTC 1723
QY 473 SerProThrProMetPhePheProLeuGlyProSerAlaProSerIleValLeuHisSerVal 492
DB 1724 AGCCCACTCCCAATGTTCTTCCCTTAAGAACATCAGAGCGGCTCAAGCTTCAAGATGTC 1783
QY 493 ThrGlyIleAspThrIleAspIleSerLeuLeuValSerSerGlyAlaValLeuSer 512

Db 1784 ACAGCATAGACACCAAGACAAAGCTTAAGACTGTGAGTCTGGGCCCAAGAAAGT 1843
QY 513 PheGluLeuLeuSerGluSerAspGlyAlaLeuMetGluHisProGluValSerGlnVal 532
Db 1844 TTGGAATTCCTCTCAAGAGACGAGTGGGCGCTTGATGAGACACCAAGAAATATCTCAAGTG 1903
QY 533 ArgArgLeuThrValGluPheLeuLeuThrAspMetProGluLeuLeuProGluAsnHisLeu 552
Db 1904 AGAGAGAAAGACTGTGAGATTAACTGACGAGATATCCAGAGATCCCGAAATATCACTCC 1963
QY 553 LysGluProLeuGluGlnSerProThrAsnLeuHisThrThrLeuLeuAspHisMetAsp 572
Db 1964 AAAAGAACCTTTGGAACATCACCAACCAATACACATACACTCAAGATCATAATGATCATGAT 2023
QY 573 ProTyrTrpAlaLeuGluAsnArgAspGluAlaHisSer 585
Db 2024 CCTTATGGGCGCTTGGAAGAACGAGATGAGACATTC 2062
RESULT 8
AF057170 2420 bp mRNA linear PRI 17-OCT-1998
LOCUS Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product,
DEFINITION complete cds.
ACCESSION AF057170.1 GI:3335160
VERSION AF057170.1 GI:3335160
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2420)
AUTHORS Petukhin, K., Koisti, M.J., Bakali, B., Li, W., Xie, G., Markneil, T.,
Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M.,
Berggren, A.A., McGarvey, Dugan, V., Figueroa, D., Austin, C.P.,
Metzger, M.L., Caskey, C.T. and Wadelius, C.
Identification of the gene responsible for Best macular dystrophy
Nat. Genet. 19 (3), 241-247 (1998)
98324772
JOURNAL
MEDLINE
PUBMED 9662395
REFERENCE 2 (bases 1 to 2420)
AUTHORS Petukhin, K.
TITLE Direct Submission
JOURNAL Submitted (03-APR-1998) Human Genetics, Merck Research
Laboratories, West Point, PA 19486, USA
FEATURES
source
1..2420
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..2420
/gene="VMD2"
/gene="VMD2"
105..1412
/gene="VMD2"
/note="Best's macular dystrophy gene; BMD; alternatively
splice isoform; contains exon 7"
/codon_start=1
/product="bestrophin"
/protein_id="AAC54344.1"
/db_xref="GI:3335161"
/translation="MTTITTSQVANAFLGSPSRLILCWRSIYKLYLGEFLILCY
IIRITVALTEEQLMPEKLTLYCDYIOLIPISFLVGFYVTLVVRMNOYENLFW
PDRMLSVSGVEKDEQGRILRLTILRYANLGNVLIRSVSTAVYRFPASOHLVOA
GFMTPAHQLEKLEKLPNNMFWVWFANLMSKAMLGRIKIDPILLOSILNMTTLR
TOGHLVADWMSISILVYTOVVAVVSFLTCLVGRQPLNPKAYGHELDLVVVP
TPIQFPYVMIVKGLSRALLGWRHGORGQOLITRMQCGQRKSRVSSQAWRT
PVLPATREAGSLEPGRKRLMWOSSSTPLEMMMLKPTGLSTGICRCPCWLWNR
CTRIGWSRTGISTSPSHSPITQLPPSSVEPLMAPSTSA"

Score: 3041.50 Matches: 582
Percent Similarity: 89.13% Conservative: 0
Best Local Similarity: 89.13% Mismatches: 3
Query Match: 97.48% Indels: 69
DB: 9 Gaps: 1
US-09-622-964A-3 (1-585) x AF057170 (1-2420)
QY 1 MetThrIleThrTyrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db 105 ATGACATCACTTACACAAAGCAAGTGGCTAATCCCGCTTAAGCTCTCTCCCGCTG 164
QY 21 LeuLeuGlyTrpArgGlySerIleTyrIleLeuLeuTyrGlyIlePheLeu 40
Db 165 CTGCTGCTGCGCGGCGAGATCTTCAAGCTGTTATGCGAGTTCTTATTTCTTG 224
QY 41 LeuGlyTyrIleIleIleArgPheIleTyrArgLeuAlaLeuThrGluGlnLeu 60
Db 225 CTGCTACTACATCATCCGCTTATTTATAGCTGCGCTCAAGCAAGCAACAGCTG 284
QY 61 MetPheGluLeuLeuThrLeuTyrCysAspSerTyrIleGlnLeuIleProIleSerPhe 80
Db 285 ATGTTAGAAAGTACTGCTGTATTGGACAGCTACATCCAGCTCATCCCATTTCTTC 344
QY 81 ValLeuGlyPheTyrValThrLeuValValThrArgTrpAsnGlnTyrGluAsnLeu 100
Db 345 GTGCTGGCTTCTACGTCGTCGTGTGTAACCGCTGTGGAAACGATCGAAGACTG 404
QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGln 120
Db 405 CCGTGGCCCAACCGCTCATGAGCTGTGTGCGGCTTCGTCGAAGCAAGCAAGCA 464
QY 121 SerArgLeuLeuArgThrLeuIleArgTyrAlaAsnLeuIleValIleLeu 140
Db 465 GCGCGCTGTGCGCGCGCACGCTATCCGTACGCAACCTGGCAACGCTCATCTG 524
QY 141 ArgSerValSerThrAlaValTyrLysArgPheProSerAlaGlnHisLeuValGlnAla 160
Db 525 CGCAGCGTACGACCGCAGTCTACAGCGCTTCCAGCGCCAGCAGCTGTGCAAGCA 584
QY 161 GlyPheMetThrProAlaGluHisLysGlnLeuGluLysLeuSerLeuProHisAsnMet 180
Db 585 GGCCTTATATGCTCCGCGACAGAACAGCTTGAGAAACTGAGCTTACACACAACTG 644
QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaThrLeuGlyValArg 200
Db 645 TTCTGGTGGCTCGGTGTGTGTGTGCAACCTGTCAATGAAGCGTGTGAGGTGCA 704
QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgThrGln 220
Db 705 ATCCGGGACCCCTATCTGCTCCAGACCTGTGAAAGAGATGAACACTTGGGTACTCG 764
QY 221 CysGlyHisLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnValVal 240
Db 765 TGTGACACCTGTATCTTCAAGCTGATTAATCCACTGGTATATACACAGGTGGTG 824
QY 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
Db 825 ACTGTGGCGGTGACAGCTTCTTCTGCACTTGTCTAGTTGGCGGCGGATTTCTGAACCA 884
QY 261 AlaAlaValTyrProGlyHisGluLeuAspLeuValValProValPheThrPheLeuGln 280
Db 885 GCCAAGGCTTACCTGTGACATGAGTGAAGTGGTGTGCTGCTTCAAGTTCTGCGAG 944
QY 281 PhePhePheTyrValGlyTyrPheLysVal 290
Db 945 TTCCTTCTATGTGTGCTGTGAGAGT-GGGCTTCCAGAGGCCCTGCTGGCTGGAG 1003
QY 290 290
Db 1004 GCATGCCAGAGGGGTATGGCAGACGCTGTTAGACAGAGATGCACTGTCAAGAAAG 1063
QY 290 290

Alignment Scores:
Pred. No.: 3.17e-260 Length: 2420

LOCUS	DEFINITION	ACCESSION	VERSION	2187 bp	mRNA	linear	PRI	26-AUG-2003
Db	1064	GAGGCTTCACGGGTTAGAAAGCAGCCAGGCGGTGGGGCCACACCTGTATTCACGTAAC	1123					
Qy	291	-----	AlaGlu	292				
Db	1124	TCGGGAGGCTGAGCGCAGAGAAATGCTTTGAACCCGGGAGCGGAGATTGTGGTGGACAG	1183					
Qy	293	GlnLeuLlIeaenProPheGlyGluAspAspAspPheGluThrAsnTrpLleValAsp	312					
Db	1184	CAGCTCAATCAACCCCTTTGGAGAGATATGATGATTTTGGACCAACACTGATGTGTGCAC	1243					
Qy	313	ArgAsnLeuGlnValSerLeuLeuLlValAspGluMetHisGlnAspGluProArgMet	332					
Db	1244	AGGAATTTGCAGGTCTCCCTGTGGCTGTGGATGAGATGCACACAGACTGCTGTGGATG	1303					
Qy	333	GluProAspMetTyTrpAsnLysProGluProGlnProProTyTrpAlaAsnAla	352					
Db	1304	GAGCGCGGACATGTACTGGAAATGAAAGCCGAGCCAGCCGCCCTTACACAGCTGTCTCCGC	1363					
Qy	353	GlnPheAspGlnValSerPheMetLysSerThrPheAsnLlSerLeuAsnLysGluGlu	372					
Db	1364	CAGTTCCTCGAGCTCTCTTTATGGGCTCCACCTTCAACATCAGCTTAACAAAGAGAG	1423					
Qy	373	MetGluPheGlnProAsnGlnGluAspGluGluAspAlaHisAlaGlyTlIeLysGlyArg	392					
Db	1424	ATGGAATTCACACCCCAATCAGAGAGACAGAGAGATGCTCAGCGTGGATCATTTGGCCGC	1488					
Qy	393	PheLeuGlyLeuGlnSerHisAspHisHisProProArgAlaAsnSerArgThrLysLeu	412					
Db	1484	TTCCTAGGCGCTGCAGATCCCATGATCAGCATCTCTCCAGGGCAAACTCAAGGCCAACTA	1543					
Qy	413	LeuTrpProLysValArgLysSerLeuLeuHisGluGluLysProLysAsnHisLysAlaAla	432					
Db	1544	CTGTGGCCCAAGAGGAAATCCCTTCTCCACAGAGGCGCTGCCCAAAACACCAAGGCAGCC	1603					
Qy	433	LysGlnAsnValLArgGlyGlnGluAspAsnLysAlaTrpLysLeuValAlaLysAla	452					
Db	1604	AAACAGAACCTTAAAGGGCCACGAGAACACAAAGAGCTTGAAGCTTGAAGCTTGAAGCTC	1663					
Qy	453	PheLysSerGlyProLeuTyTrpGlnArgProGlyTyTrpSerAlaProGlnThrProLeu	472					
Db	1664	TTCAAGTCTGCCCCCACTGTATTCAGAGCGCAGGCTCTCAAGTGGCCCCACAGCGCCCTTC	1723					
Qy	473	SerProThrProMetPhePheProLeuGlnProSerAlaProSerLysLeuHisSerAla	492					
Db	1724	AGCCCACTCCCAATGTTCTTCCCTTAAGAACATGAGCGCGTCAAACTTCACACGTGTC	1783					
Qy	493	ThrGlyLlIeaenThrLysAspLysSerLeuLysThrValSerSerGlyAlaLysLysSer	512					
Db	1784	ACAGGCAATGACACCAAAAGCAAAAGCTTAAAGATGTGATGTTGGGGCCAAAGAAAGT	1843					
Qy	513	PheGluLeuLeuSerGluSerLysThrAspGlyAlaLeuMetGluHisProGluValSerGlnVal	532					
Db	1844	TTTGAATATGCTCTCAGAGAGCGATGGGGCTTTGATGAGGACGCCAGAAAGTATCTCAATGT	1903					
Qy	533	ArgArgLysThrValGluPheAsnLeuThrAspMetProGluLlIerProGluAsnHisLysLeu	552					
Db	1904	AGAGAGAAACTGTGGAAGTTTAACTTGACGGATATGCCAGAGATCCCGGAAAATCACTTC	1963					
Qy	553	LysGluProLeuGluGlnSerProThrAsnLlIeHisThrThrLysLysAspHisLysMetAsp	572					
Db	1964	AAAGAACTTTGGAAACATTCACCAACCAACATCACTACACTAAAGATCACTGGAT	2023					
Qy	573	ProTyTrpTrpAlaLeuGluAsnArgAspGluLalaHisSer	585					
Db	2024	CCTTATTTGGGCTTGGAAAACAGGAGATGAAGACATTTCC	2062					

KEYWORDS	
SOURCE	Macaca fascicularis (crab-eating macaque)
ORGANISM	Macaca fascicularis
	Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; Macaca.
REFERENCE	1 (bases 1 to 2187)
AUTHORS	Okamoto,H., Umeda,S., Suzuki,M.T., Yoshikawa,Y., Tanaka,Y. and Iwata,T.
TITLE	Molecular Cloning of VMD2 Gene from Cynomolgus Monkey (Macaca fascicularis)
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 2187)
AUTHORS	Okamoto,H., Umeda,S., Suzuki,M.T., Yoshikawa,Y., Tanaka,Y. and Iwata,T.
TITLE	Direct Submission
JOURNAL	Submitted (01-AUG-2003) National Tokyo Medical Center, National Institute of Sensory Organs, 2-5-1, Higashtgaoka, Meguro-ku, Tokyo 152-0021, Japan
FEATURES	Location/Qualifiers
source	1..2187
	/organism="Macaca fascicularis"
	/mol_type="mRNA"
	/db_xref="taxon:9541"
	1..2187
	/gene="VMD2"
	67..1824
	/gene="VMD2"
	/codon_start=1
	/product="bestrophin"
	/protein_id="AA056049.1"
	/db_xref="GI:34013783"
	/translation="MTITYSQVANAALGSPSRLLLCMRGSYIKLYLGEFFIFLLCYV IIRYIRALTEBOQLMFEKLYLDCSYIOLIPISFGLPVYTLVVRMNOYENLP PDRPLSVSGPVEGKDCOGRLRPRLIRYANLGNVLTREYSTAVYRPFSAOHLNOA GMPDPAHKKQLEKSLPNNFWPWFPANLSKAMLGIGIRPILLOSILNENNTLR TQCHLAIDYDWISILPLYTVVTVAVYSPFLTCLVGRQINPAKAYGHELDLVVAF TFQLPFYVGMCLKVAEQLINPFGEDDDPFTNMIVDRNLGVSLAVDEMQDLPRMEP DMVNEDEPHRYPAASOPRASFMGSTFNI SLNKEMFQPOQEDKDAHTGIER FLIGASHDHHIPGANSRRTKLMPRESLTHEGLVQKHKAVKONVGRGLDKNAWLKAV DAKRSAPLYORPGYNSAPQTPLSPTPMPFPPEPVSLSKLSHYDTPDKSLKTVSSG AKNSSELSSGDALMEHPEVSHVRKTRVENLDMPEIRFENHLKEPLELSTNIHAT LKQHDVPMALENRDEAHS"
ORIGIN	
Alignment Scores:	
Pred. No.:	1.11e-256
Score:	3001.00
Percent Similarity:	96.92%
Best Local Similarity:	96.41%
Query Match:	96.19%
DB:	9
	Gaps: 0
US-09-622-964A-3 (1-585) x AY357925 (1-2187)	
Oy	1 MetThrlleThrTyTrhSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db	67 ATGACCATCACTACAGACCAAGTGGCTAATCCCGCTTAGAGCTCTTCCGGCTG 126
Oy	21 LeuLeuCystrPaArgIySerIleTyTrLysLeuLeuTyrgIyGluPheLeuIlePheLeu 40
Db	127 CTGCTGAGCTGGCGGGGACAGCATCTACAGGCTGCTGATG6GCGAGTCTTCATCTTCTG 186
Oy	41 LeuCystrTyrrIleIleIeArgPheIleTyArgLeuAlaLeuThrGluGluGlnLeu 60
Db	187 CTCGCTACCACTACATCCGCTTATTATTATTAAGCTGGCCCTCACAAGAGCAACACTG 246
Oy	61 MetPheGluIyrsLeuThrLeuTyCyAspSerTyrrIleGlnLeuIlePheIleSerPhe 80
Db	247 ATGTTTAGAAGCTGACTCTGTATTGGGACAGCTACATCAGCTCATCCCATTTCTTC 306
Oy	81 ValIeuGlyPheTyValIThrLeuValIThrArgTrpTrpAengIlyrGluAsnLeu 100

Db 307 GTGCTGGGCTTCTAGTGAACGCTGGTCTGACCCGCGTGGTGGACCAAGTGAAGAACCTG 366
 Qy 101 PCTTTPCTAAspArgLeuMetSerLeuValSerGlyPheValGluGlyLeuAspGluGln 120
 Db 367 CCGTGGCCGACGCGCTCAATGAGCTGGTGGGCTTCTGTTGGAAGGACCAAGACCAAG 426
 Qy 121 SerArgLeuLeuArgThrLeuIleArgTyrAlaAsnLeuGlyAsnValLeuIleLeu 140
 Db 427 GCGCGGCTGCTGGCGGACGCTCAATCCGCTACGCAACCTGGGACAGCTGCTCATCTTG 486
 Qy 141 ArgSerValSerThrAlaValTyrLeuArgPheProSerAlaGlnIleLeuValGlnAla 160
 Db 487 GCGACGCTGACGACGCGCTCAAGGCGCTTCCGACGCGCCACGACCTGGTGAACA 546
 Qy 161 GlyPheMetThrProAlaGluHisGlySerLeuGlyLeuSerLeuProHisAsnMet 180
 Db 547 GCGTTTATGACCCCGGACAGAACAGAGTGGAGAAAGCTGAGTTTACACACAAACATG 606
 Qy 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetLysAlaTrpLeuGlyArg 200
 Db 607 TTTCTGGGTGCGCTGGGTGTGGTTGCCAACCTGTCAATGAAAGCGTGGCTTGGAGTCA 666
 Qy 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgThrGln 220
 Db 667 ATCCGGGACCCCTATCCCTGCTCCAGAGCCTGGCTGAACGAGATGAACACCTTGGCAG 726
 Qy 221 CysGlyHisLeuTyrAlaTyrAspTrpIleSerIleProLeuValTyrThrGlnVal 240
 Db 727 TGTGACACCTGTATGCTTACGACTGAGATTGATCCACGCTGGTGTATACACAGGTGTG 786
 Qy 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
 Db 787 ACTGGGGGGGTGATCAGCTTCTTCTGACTTGTCTAGTTGGCGGAGTTTCTGACCCA 846
 Qy 261 AlaLysAlaTyrProGlyHisGlnLeuAspLeuValValProValPheThrPheLeuGln 280
 Db 847 GCGAAGGCTACCCCTGCGCATGAGCTGAGCTGCTGTGGCCGCTTCCACATTCCTGAG 906
 Qy 281 PhePhePheTyrValGlyTrpLeuLeuValAlaGluGlnLeuIleAsnProPheGlyGlu 300
 Db 907 TTTCTTCTTATGTTGTGGTGTGAAGGTGGCAGACGCTCATCAACCCGTTTGGAGAG 966
 Qy 301 AspAspAspAspPheGluThrAsnTrpIleValAspArgLeuLeuValSerLeuLeu 320
 Db 967 GACCATGATGATTTTGGACCACTGATGTCCAGAGAAATTGGCAGGTCTCCCTGCTG 1026
 Qy 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrTrpAsnLys 340
 Db 1027 GCTGTGATGATGATGACACAGGACCTGCTGATGAGGCGACACATGTACTGGAATGAG 1086
 Qy 341 ProGluProGlnProProTyrThrAlaAlaSerAlaGlnPheArgTGAAspPheMet 360
 Db 1087 CTTGAGGACACACCCCTTACAGAGCTGCTTCCGCCCAAGTTCCGTCAAGCTCTTTATG 1146
 Qy 361 GlySerThrPheAsnIleSerLeuAsnLysGluGluMetGluPheGlnProAsnGlnGlu 380
 Db 1147 GGCCTCCACCTTCAACATCAGCTGACCAAGAGAGAGATGAGTTTCAGCCCAATCAGAG 1206
 Qy 381 AspGluGluAspAlaHisAlaGlyTyrIleGlyArgPheLeuGlyLeuGlnSerHisAsp 400
 Db 1207 GACAAAGAGAGCGCTCACTGCGCATATGGCGCTTCCATGAGGCTGCGAGTCCCATGAT 1266
 Qy 401 HisHisProProArgAlaAsnSerArgThrLysLeuLeuTrpProLysArgLysLeu 420
 Db 1267 CACCATCTCTCCGGGGGCAAACTCAAGAACCAACTACTGTGGCCCAAGAGGATCCCTT 1326
 Qy 421 LeuHisGluGlyLeuProLysAsnHisLysAlaAlaLysGlnAsnValArgGlyGlnGlu 440
 Db 1327 CTCACAGAGGCGCTGCCCAAAACCAACAGGCGCTCAAAACAGAAAGCTTGAAGGCGTGA 1386
 Qy 441 AspAsnLysAlaTrpLysLeuLysAlaValAspAlaPheLysSerGlyProLeuTyrGln 460
 Db 1387 GACAAAGAGGCGCTGAAAGCTTAAAGCTGTAGATGCTTCAAGTCTGCCCACTGTATCAG 1446

Qy 461 ArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrProMetPhePhePro 480
 Db 1447 AGGCGAGGCTTACTACAGATGGCCCCACAGACGCCCTCTCAGCCCTCCATCGATTTCTTCCC 1506
 Qy 481 LeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAspTrpLysAspLys 500
 Db 1507 CCAAGACCATCAAGCTGCTGTAACAGCTTCAAGTGTCAAGGCAATGACACCAAGACAA 1566
 Qy 501 SerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeuLeuSerGluSerAsp 520
 Db 1567 AGCTTAAAGACTGAGTTCCTGGGGCCCAAGAACGTTCTGAATTCCTCAGGAGGAGAT 1626
 Qy 521 GylAlaLeuMetGluHisProGluValSerGlnValArgArgLysThrValGluPheAsn 540
 Db 1627 GGGGCTTGATGAGACACCCAGAAATCTCACAGTGAAGGAGAAACCCGTGAGTTTAA 1686
 Qy 541 LeuThrAspMetProGluLeuProGluAsnHisLeuLysGluProLeuGluGlnSerPro 560
 Db 1687 CTGACGATATGCCAGAGATCCCGAAATCATCTCAAGAACCTTTGGAACTATCAACA 1746
 Qy 561 ThrAsnIleHisThrThrLeuLysAspHisMetAspProTyrTrpAlaLeuGluAsnArg 580
 Db 1747 ACCAATCATCATGCTCACTAAAGATCACTGATCTTATTGGGCTTGAAGAAACAGG 1806
 Qy 581 AspGluAlaHisSer 585
 Db 1807 GATGAAGCACATTC 1821

RESULT 10
 BC041664
 LOCUS
 DEFINITION
 Homo sapiens vitelliform macular dystrophy (Best disease, bestrophin), mRNA (cDNA clone MGC:47884 IMAGE:5194649), complete cds.
 BC041664
 BC041664.1 GI:27371319
 VERSION
 KEYWORDS
 MGC.
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2435)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Valek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schultz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E., Schnerk, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 MEDLINE
 PUBMED
 12477932
 2 (bases 1 to 2435)
 Strausberg, R.
 DIRECT SUBMISSION
 TITLE
 JOURNAL
 Submitted (20-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nigr.nih.gov
Ahter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooke,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Hachiguchi,P.,
Hansen,N., Ho,S.-L., Karlins,R., Kwong,P., Latic,P., Legaspi,R.,
Mazuro,Q.L., Masiello,C., Maseri,B., Mastrian,S.D., McCloskey,J.C.,
McGowan,J., Pearson,R., Stancipod,S., Thomas,P.J., Touchman,J.W.,
Taurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNLT at: <http://image.llnl.gov>
Series: IRAK Plate: 82 Row: d Column: 4.
Location/Qualifiers

FEATURES
source

1. .2435
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:47884 IMAGE:5194649"
/issue_type="Brain, adult, 6 pooled whole brains"
/clone_id="NIH_MGC_114"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"

gene
1. .2435
/gene="VMD2"
/note="synonyms: BMD, BEST"
/db_xref="locusid:7439"
/db_xref="MIM:153700"
87.11901
/gene="VMD2"
/codon_start=1
/product="VMD2 protein"
/protein_id="AAH41664.1"
/db_xref="GI:2371320"
/db_xref="locusid:7439"
/db_xref="MIM:153700"
/translation="MFEKLTLYCDSTYQILIPISFLGFFVTLLVTRMMNOYENLPWPD
RLMSLVGFVGEKDEQGLRLRLIRYANLGNVLIRSVTAVKRFPASQHLVQAGF
MTPAEHKOLEKLSLPHNMFVPMVPMANLSKAWLGGIRIDPILLOSLNEMTLRTO
CGHLXADWTSLPLVYQVTVAVYSPFLCTVGRPLNPKAKPGHELDLVVPTPL
LQFFPYGWLKVAEQLINPGEDEDDDEFTTNKTVDRNLQVSLAVDENHODLPMEPDM
YWNKPEOPPYTAASQFRVRSFGSTFNISLKEEMFQPNQDESDALAGIGRL
GLQSHDHPPLANSRTKLMPKRESLLHEGLPKNHKAKQVNRQEDNKWLAVNA
FKSAPLYORPGYAPQPLSPMPFPLPSAPSKLSVTGIDTKDSLKTVSGAK
KSPFLSESDGALMEHPESQVRKRTVEFNTDMPRIPEMLKELPSPPNHTTK
DHMPRYALENRSVLHNOGHCIACRPASLALPRLNHLGPHNCGSTLDLRPAL
AMGILYLAFTGILKCSGSPFLSPMHPEDFLGPEGR"

misc_feature

90. .1022
/gene="VMD2"
/note="DUF289; Region: Putative membrane protein. This
family called family 8 in, may be a transmembrane protein
The specific function of this protein is unknown"
/db_xref="CDD:pfam01062"

ORIGIN

Alignment Scores:

Pred. No.: 2.94e-240 Length: 2435
Score: 2817.00 Matches: 527
Percent Similarity: 98.87% Conservative: 0
Best Local Similarity: 98.87% Mismatches: 6
Query Match: 90.29% Indels: 0
DB: 9 Gaps: 0

US-09-622-964A-3 (1-585) x BC041664 (1-2435)

Qy	52	LeuAlaLeuPhrGluGluGlnLeuMetPheGluLeuThrLeuTyCysAspSer	71
Db	60	CTGGCCCTCAGGAAACACACGTATGTTTAAAGAACTGTGATTTGACACGC	119
Qy	72	TyrIleGlnLeuIlePheIleSerPheValLeuGluPheTyValThrLeuValThr	91
Db	120	TACATCAGCTCATCCCATTTCTCTGTGTGGCTTACAGTACGCTGTGTGACC	179
Qy	92	ArgTrpTrpAsnGlnTyGluAsnLeuProTrpProAspArgLeuMetSerLeuValSer	111
Db	180	CGCTGGTGAACCGATCGAGAACCTCGGCGCCGACCGCTCATGAGCCTGTGTGCG	239
Qy	112	GlyPheValGluGlyValAspGluGlnSerArgLeuLeuArgTrpLeuIleArgTyr	131
Db	240	GGCTTCCTCAAGGCAAGACAGACAAAGCGGCTGTGGCGGACGCTCATCGGCTAC	299
Qy	132	AlaAsnLeuGlyAsnValLeuIleLeuArgSerValSerThrAlaValTyTyrAspPhe	151
Db	300	GCCAACTGGGCAACGTCCTCATCTCGACGCTCAGACCGGACGCTACAGCGCTTC	359
Qy	152	ProSerAlaGlnIleLeuValGlnAlaGlyPheMetThrProAlaGluHISLeuGlnLeu	171
Db	360	CCCAAGCCGCGACACTGTGTGCAAGCAGGCTTTATGACTCCGGCAGAACACACAGCTTG	419
Qy	172	GluIleLeuSerLeuProHisAsnMetPheTrpValProTrpValITrPheAlaAsnLeu	191
Db	420	GAGAACTGAGCCCTACACACACACATGTTCTGGGTGCGCTGGGTGGTTGGCCAACTGG	479
Qy	192	SerMetIleAlaTrpLeuGlyGlyArgIleArgAspProIleLeuLeuGlnSerLeuLeu	211
Db	480	TCAATGAGCGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG	539
Qy	212	AsnGluMetAsnThrLeuArgTrpGlnGlyAsnLeuTyAlaTyAspTrpIleSer	231
Db	540	AACGAGATGAACACCTTGCGTACAGTGTGACACCTGTGCTGCTGCTGCTGCTGCTG	599
Qy	232	IleProLeuValTyThrGlnValValThrValAlaValTySerPhePheLeuThrCys	251
Db	600	ATCCACGTGGTATACACAGTGTGATCTGTGGGTGTGATCTGTGTGTCTTCTTCTTCTG	659
Qy	252	LeuValGlyArgGlnPheLeuAsnProAlaValAlaTyProGluIleGluLeuAspLeu	271
Db	660	CTAGTGGCGGCGGCTTCTGTAACCCAGCCAGCCCTTACCTGGCCATGAGTGCCTC	719
Qy	272	ValValProValPheThrPheLeuGlnPhePheTyValGlyTrpLeuValAla	291
Db	720	GTGTGCGCGCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	779
Qy	292	GluGlnLeuIleAsnProPheGlyGluAspAspAspPheGluThrAsnTrpIleVal	311
Db	780	GAGAGCTCATCAACCCCTTGGAGAGATGATATATTTTGAGACCACTGATATGTC	839
Qy	312	AspArgAsnLeuGlnValSerLeuLeuAlaValAspGluMetHISGlnAspLeuProArg	331
Db	840	GACAGAAATTTGACGGTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	899
Qy	332	MetGluProAspMetTyTrpAsnIleAspGluProGluProProTyTyThrAlaAlaSer	351
Db	900	ATGAGGCGGACATGTATGTAATTAACCCGAGCAGCCGCCCTTACAGCTGCTTCC	959
Qy	352	AlaGlnPheArgArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnGlu	371
Db	960	GCCAGTTCGTGAGTCTCTTATGCGCTCCACTTCAACATCAGCTGAAACAAAGAG	1019
Qy	372	GluMetGluPheGlnProAsnGlnIleAspGluIleAspAlaHISAlaGlyIleIleGly	391
Db	1020	GAGATGAGTTCAGGCCCATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1079
Qy	392	ArgPheLeuGlyLeuGlnSerHisAspHisHisAspProArgAlaAsnSerArgThrIle	411
Db	1080	CGTTCTTACGAGCTGAGTCCCATGATCACCATCTCTCCAGGCGAACTCAAGAGACCAA	1139


```

QY 412 LeuLeuTPProLySarGluSerLeuLeuHISGluGlyLeuProLySaAenHISlyAla 431
DB 1140 CTACGTGGCCCAAGAGGGAATCCCTTCCACAGAGGCTGGCCCAAAAACCAAGAGCA 1199
QY 432 AAlaySGIaenValaArglyGInGluAspAsnLySaLaTPrLySLeuLySaLaValaAsp 451
DB 1200 GCCAAACAGAACTTAAAGGGCCAGAGAGACAAACAGAGCTTGAAGCTTAAAGCTGTGGAC 1259
QY 452 AAlaPheLySergLyProLeuTyTrGlnaArgProGlyTyTrTyTrSerAlaProGlnTrPro 471
DB 1260 GCCTTCAAGTCTGCCCACTGATTCAGAGGCAGGCTACTAAGTCCAGCCCAAGAGAGCCCC 1319
QY 472 LeuSerProThrProMetPhePhePheProLeuGluProSerAlaProSerLySLeuHISsar 491
DB 1320 CTCAGCCCACTCCCACTGCTTCCCTCCAGAACATCAGCCGCCCTCAAGCTTCAAGT 1379
QY 492 ValThrGlyIleAspThrLySaAspLySergLySLeuLyThrValSergLyAlaLySlyS 511
DB 1380 GTCAAGGCAATAGACACAAAGCAAAAGCTTAAAGACTGTGAGTTCTGGGCGCAAAA 1439
QY 512 SerPheGluLeuLeuSerGluSerGluSerAlaAlaLeuMetGluHISProGluValSergIn 531
DB 1440 AGTTTGAATTGCTCTCAGAGACCATGGGCTTATGAGAGACCCAGAGATATCTCA 1499
QY 532 ValArgArgLyThrValGluPheAsnLeuThrAspMetProGluIleProGluAsnHIS 551
DB 1500 GTAGAGAGGAAACTGTGAGTTAACTGACGAGATATCCAGAGATCCCGAAATATCAC 1559
QY 552 LeuLyGluProLeuGluGlnSerProThrAsnIleHISThrThleuLySaAspHISmet 571
DB 1560 CTCAAAGAACTTTGGAACATCACCAACCAATACATACACATACACATCAAGATCACTG 1619
QY 572 AspProTyTrTPAlaLeuGluAsnArgAspGluAlaHIS 584
DB 1620 GATCCTTATTTGGCTTGGAAAAAGTGTCTGCTCCAC 1658

```

RESULT 11
BC015220
LOCUS Homo sapiens vitelliform macular dystrophy (Best disease, bestrophin), mRNA (CDNA clone MGC:17761 IMAGE:3877806), complete cds.
DEFINITION BC015220 2441 bp mRNA linear PRI 07-OCT-2003
ACCESSION BC015220
VERSION BC015220.2 GI:34783515
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Strausberg, R.L., Feingold, R.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heile, F., Ditchen, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stedelson, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wolley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Blakeley, R.W., Touchman, J.W., Green, E.D., Bouffard, G.G., Rodriguez, A.C., Gilmwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Schmech, A., Schein, J.E., Jones, S.J., and Marra, M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257

FEATURES
SOURCE
gene
CDS
misc_feature
ORIGIN

REFERENCE PUBMED 12477932
AUTHORS 2 (bases 1 to 2441)
TITLE Strausberg, R.
JOURNAL Direct Submission
 Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Sep 16, 2003 this sequence version replaced gi:21955361.
 Contact: MGC help desk
 Email: cgapsb-remail.nih.gov
 Tissue Procurement: DCTD/DTF/Gazdar
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www.ahgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Gilmwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 14 Row: b Column: 9.

FEATURES
SOURCE
gene
CDS
misc_feature
ORIGIN

1..2441
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:17761 IMAGE:3877806"
 /tissue_type="lung, large cell carcinoma"
 /clone_lib="NIH_MGC_68"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 1..2441
 /gene="VMD2"
 /note="synonym: BMD, BEST"
 /db_xref="locusid:7439"
 /db_xref="MIM:153700"
 573..2069
 /gene="VMD2"
 /codon_start=1
 /product="VMD2 protein"
 /protein_id="AAH15220.1"
 /db_xref="gi:22800661"
 /db_xref="locusid:7439"
 /db_xref="MIM:153700"
 /translation="MFEKLTLYCDSTYLTLPISFVLGFTLVVTRMNOYENLPMPDRLMSLVSGFVGKDEQGLRLRLIRYANGVILRSVSTAVYKRPSPASQHLVOAGFMTPEAHKQLEKLSLPHNMFVWVPMVFNLSKMLIGRIDPILLOSILNENNTLPTGCHLYAVDMSIPLVTVQVTVAVASFLICVIGROINPAKAYRGEHLDVVPVTFEPLNLSNEMEFQNOEDBEDAHGILGRFLGQSODHDPNANSRTKLMPRESLHGGIKPKHAKQNVROQEDNKAKWAKVADAFKSAPLYQRPQYVSAFQTPISPMFPLEPSAPSKHAKVQIDTKDSILKTVSSGAKSEFELISEDGALMEHPEVSGVRRRTVEFNLTDPBEIPENILKEPLRQSPNITHTLKHMDPYMALENDEAHS"
 576..1427
 /gene="VMD2"
 /note="ndup289, Region: Putative membrane protein. This family called family 8 in, may be a transmembrane protein The specific function of this protein is unknown"
 /db_xref="CDD:pfam01062"

ALIGNMENT SCORES
 Pred. No.: 2 09e-227
 Score: 2672.50
 Percent Similarity: 94.57%
 Best Local Similarity: 94.57%
 Query Match: 85.66%
 DB: 9
 Gaps: 1

US-09-622-964A-3 (1-585) x BC015220 (1-2441)

OY	52	LeuAlaLeuThrGluGluGlnGlnLeuMetPheGlyIleuThrLeuTyrCYAspSer	71
Db	546	CTGGCCCTCAGCGAAGAACACAGCTGATGTTTGAAGAACTGACTCTGATATGGCCAGAC	605
OY	72	TyrIleGlnLeuIleProIleSerPheValLeuGlyPheTyrValThrLeuValAlaThr	91
Db	666	TACATCCAGCTCATCTCCATTTCTTCGTGCTGGGCTTCTAGTACACCTGGTCTGCAC	665
OY	92	ArgTyrTPraenGlnTyrGluAsnLeuProTPraeAPargLeuMetSerIleuValSer	111
Db	666	CGCTGGTGAACCAACTAGACGAACCTGGCCGTGGCCGACCGGCTCATGAGCTCGTGTGCG	725
OY	112	GlyPheValGluGlyAspGlnIleSerArgLeuLeuArgTyrThrLeuIleArgTyr	131
Db	726	GAGCTTCGTGAAAGGGAAGGACGAGCAAGCCGAGCTGTGGCGGAGCGTCACTCCGCTAC	785
OY	132	AlaAsnLeuGlyAsnValLeuIleLeuArgSerValSerThrAlaValTyrIleAspPhe	151
Db	786	GCCAACTGGGGCAACTGTCTCATCTGTGGCAGCGTCAGACCGGAGCTGTCAACAGCGCTTC	845
OY	152	ProSerAlaGlnIleLeuValGlnAlaGlyPheMetThrProAlaGluIleIleGlnLeu	171
Db	846	CCCAACGGCCACAGCACTGTGTGCACAGACGCTTTAAGTCTCCGGCAGAAACACAGCAAGTTCG	905
OY	172	GluIleLeuSerLeuProHisAsnMetPheTyrValProTyrValTyrPheAlaAsnLeu	191
Db	966	GAGAAACTGAGGCTTACACACACACTGTTCTGGGTGGCCCTGGGTGTGGTTGCCAACCTG	965
OY	192	SerMetCysAlaTyrLeuGlyGlyValArgIleArgAspProIleLeuLeuGlnSerLeuLeu	211
Db	966	TCAATGAAGAGCGTGGCTTGGAGGTGGAATCCGGGACCCCTATCTCTGCTCCAGAGCGCTGCTG	1025
OY	212	AsnGluMetAsnThrLeuArgThrGlnCysGlyGlnIleuTyrAlaTyrAspTyrIleSer	231
Db	1026	AACGAGATGAAACACTTGGCTACTAGTGTGGACACTGTATGCTTACAGACTGGATTAGT	1085
OY	232	IleProLeuValTyrThrGlnValIleThrValAlaValTyrSerPhePheLeuThrCys	251
Db	1086	ATCCCACTGGTGTATACACAGGTGTGACTGTGGCGGATGACAGCTTCTTCTCGACTTGT	1145
OY	252	LeuValGlyIleArgGlnPheLeuAsnProAlaIleValIleTyrProGlyHisGluLeuAspLeu	271
Db	1146	CTAGTTGGGCGGCACTTCTCTGAACCCAGCCAGAGGCTTACCTCGGCAAGAGCTGGACCTC	1205
OY	272	ValValProValPheThrPheLeuGlnPhePhePheTyrValGlyTyrLeuIleValAla	291
Db	1206	GTTGGCCCGCTTCAAGTCTCTGCAAGTCTTCTTCTTATGTTGGCTGACTGAAG-----	1255
OY	292	GluGlnLeuIleAsnProPheGlyIleuAspAspAspPheGluThrAsnTyrIleValAl	311
Db	1259	-----	1259
OY	312	AspArgAsnLeuGlnValSerIleuLeuAlaValAspGluMetHisGlnAspLeuProArg	331
Db	1260	-----GATGCCCTGTGGCTGTGGATGAGATGACACCAAGACCTGCTGG	1304
OY	332	MetGluProAspMetTyrTyrAsnIleuAspProGluProGlnProProTyrThrAlaIleSer	351
Db	1305	ATGAGGCCGAGACATGATGATCGGAATAAAGCCGAGCCAGACAGCCCTTACACAGCTCTTCC	1364
OY	352	AlaGlnPheIleArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnIleuGlu	371
Db	1365	GCCCAAGTTCGTCGAGCGCTCTTATGGAGCTTCACACTTCAACATGAGCTGAACAAAG	1424
OY	372	GluMetGluPheGlnProAsnGlnIleuArgGluGluAspAlaHisAlaGlyIleIleGly	391
Db	1425	GAGATGGAGTTCCAGCCCAATGAGGAGGACAGAGAGATGCTCAAGCTGGCATCTTGGC	1484
OY	392	ArgPheLeuGlyLeuGlnSerHisValAsnHisIleProTyrAlaAsnSerArgThrIleu	411
Db	1485	CGCTTCTTAGGCTCTGACATGCTCCATATGACCACTTCTCCAGGGCCAACTCAAGAGCCAA	1544

OY		412	Leu1eUtrPrProLyVaRgGlUsErLeuLeuH1seL1uLyLeuPrOlyVaEnH1sLyA1a	431
OY		412	Leu1eUtrPrProLyVaRgGlUsErLeuLeuH1seL1uLyLeuPrOlyVaEnH1sLyA1a	431
Db		1545	CTACTGTGGCCCAAGAGGAATCCCTTCTCACAAGGACCGGCCCAAACACAAAGGCA	1604
OY		432	AlaLySeGiAnPaValArGSLyGLNGlAbApRaMnLyVaLaTRPLySLyALaValAPr	451
Db		1605	GCCAAACAGAACCATTAGGGGCCAGAGAACACAAAGCCTTGAAAGCTTAGGCTGTGCAC	1664
OY		452	AlaPheLySerGiLyProLeuTyTrGLNaRProGiLyTyTyTSerAlaProGiInThrPro	471
Db		1665	GCTTTCAAGTCGTCCCCCACTGTATCACAGGSCAGGCTACTACAGTGCCTCCCAAGACACC	1724
OY		472	LeuSerProThrPrPrometPhePheProLeuGlnProSerAlaProSerLyLeuH1sSer	491
Db		1725	CTCAGCCCCCATCCCATGTTCTTCTCCCTTAAGAACCATCAGGCGCGTCAAAAGCTTCAAGT	1784
OY		492	ValThrGylIleAspThrLyAspLySeserLeuLyThrValIseserGylVALaLyLyS	511
Db		1785	GTCAcAGGCATAGAcACCAAGAACAAAAGCTTAAAGACTGTGAAGTTCTGGGSCCAADAA	1844
OY		512	SerPheGluLeuLeuSerGiLySerAspGylVALaLeuWetGluH1sProGluValSerGln	531
Db		1845	AGTTTGAATGCTCTCAGAGAGCGAGTGGGGCTTGTATGGAGCACCCAGAAAGTATCTCAA	1904
OY		532	ValArgaGdyLySThrValGluPhaAntLeuThyrMetProGiLiLeProGluVaEnH1s	551
Db		1905	GTGAGAGAGAAACCTGTGGAGTTTAACTGACGGAATATGCCAGAGATCCCGGAAAAATCAC	1964
OY		552	LeuLySeGiUProLeuGluGlnSerProThrAmLiLeH1sThrThrLeuLyAspH1sMet	571
Db		1965	CTCAAGAAGCTTTGGAAACATCAACCAACCAACTACACTACACTCAAAAGATCACATG	2024
OY		572	AapProTyTrTPdAlaLeuGluAenArgAspGluAhaH1ser	585
Db		2025	GATCCTTAATGGCGCTTGAAAAACAGGATGAAACACATTC	2066
RESULT 12				
BC079048				
LOCUS	BC079048	2035 bp	mRNA	linear ROD 03-AUG-2004
DEFINITION	Rattus norvegicus cDNA clone MGC:93985 IMAGE:7115295, complete cds.			
ACCESSION	BC079048			
VERSION	BC079048.1	GI:50927692		
KEYWORDS	MGC.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
REFERENCE	1 (bases 1 to 2035)			
AUTHORS	Strausberg, R.L., Fellingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Bueltow, K.H., Schaefer, C.F., Bhac, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hale, F., Datchenko, L., Marusina, K., Farmer, A.O., Rubin, G.M., Hong, L., Stadlton, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L., Schneez, T.E., Brownstein, M.J., Uddin, T.B., Tohyuki, S., Caminci, P., Prange, C., Kaha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Boesk, S.A., McEwan, P.J., McNernan, K.C., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Ketteman, M., Medan, A.C., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonifard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimm, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, J.I., Skalska, U., Smalys, D.E., Bucherich, A., Schein, J.E., Jones, S.J. and Marx, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (26),	16899-16903	(2002)
PUBMED	12477932			
REFERENCE	2 (bases 1 to 2035)			

AUTHORS

Director MGC Project.
Direct Submission
Submitted (02-AUG-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Howard Jacobs
cDNA Library Preparation: Express Genomics
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LNL)
Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcddpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

FEATURES

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 183 Row: a Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis. Similarity but not identity to protein.

SOURCE

1. 2035
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:93985 IMAGE:7115295"
/tissue_type="Testis, Rat (Brown Norway)"
/clone_lib="MGC_237"
/lab_host="DH10B"
/note="Vector: pXpress1"
153. .1805
/product="Unknown (protein for MGC:93985)"
/protein_id="AAH79048.1"
/db_xref="GI:50927693"
/translation="MTITVTNANARLGSFSCLLLRNGSYIKLLYGEPLVFLFY
CIRGLRYNLNHHQOQFEKLLYCNSTYIPIPISEFVYVTVLTVSRMNOYENLMA
PRLNMYSSFVGEKQEQEGLRLRLTLYALIGVLLIRSTVYKRPBLOHLMQ
GMPTRGEYKQRLGLPHNTEFVPMWPMANLAKYIGGRIRDRVILLQSLMREYCTIK
TQCGHLYADWLNIPLYTVYVAVSFFLACILGRPLNPLKNTYGEHMDLVPEP
TLIOFLVGMKLKVAEOLINPGEHDDPEFTNWTIDRLQVLSLVGMHDDVPEP
DLYNBAAPORPYTAASAGSRQSFMTSLKKDSEESTYDVKKENDYGSITG
PRIGQSDYHPKTDLTKLKMSKQBFLEDSKNDKOKDKNDVWRPDLFRALPK
KRBPYRQAPQLSHLPMVPPPOQASAPSVAYSDTGGGSPSAQYEAQCHMKKTVEFNL
NIPCTPHEPQPHLDQMTIILKEBESYFVSDKADTPVSYK"

CDS

Alignment Scores:
Pred. No.: 4,63e-160 Length: 2035
Score: 1914.00 Matches: 373
Percent Similarity: 74.91% Conservative: 63
Best Local Similarity: 64.09% Mismatches: 106
Query Match: 61.35% Indels: 40
DB: 10 Gaps: 7

ORIGIN

US-09-622-964A-3 (1-585) x BC079048 (1-2035)

QY 1 MetThrlleThrYrThrSerGlnValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
DB 153 ATGACTATACACTACACAAACGAAGTGGCCAAATGCCGCTTGTTCTTCTGTCGCTC 212
QY 21 LeuLeuGlySerGlySerIleYrYrLeuLeuYrGlyGluPheLeuIlePheLeu 40
DB 213 CTCTCGGCGCGGAGGACGACATCTACAGCTCTCTACGGAATTCCTTGCTTCAATA 272
QY 41 LeuGlyYrYrIleIleArgPheIleYrYrLeuAlaLeuThrGluGluGlnGlnLeu 60
DB 273 TTCCTTACTATATGACATCCGTGACATCTACAGAAATGTTCTCTCAATATATGACGAC 332

QY 61 MetPheGluYrLeuThrLeuYrCysAspSerYrIleGlnLeuIleProIleSerPhe 80
DB 333 CAGTTTGAAAGTGGCTCTGTACTGCAACAGCTACATTCAGTTCAATCCCTATCTCTC 392
QY 81 ValLeuGlyPheYrValThrLeuValValThrArgTrpPheGlnIleYrGluAsnLeu 100
DB 393 GTTCTGGGTTCTTANTGATGATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 452
QY 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyYrAspGluGln 120
DB 453 CCGTGGCCGACCCCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 512
QY 121 SerArgLeuLeuAArgTrpLeuIleArgYrAlaAsnLeuGlyAsnValLeuIleLeu 140
DB 513 GCGCGGTTGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 572
QY 141 ArgSerValSerThrAlaValYrYrLeuArgPheProSerAlaGlnIleSerValGlnAla 160
DB 573 GCGAGCATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 632
QY 161 GlyPheMetThrProIleGluIleGlyGlnLeuGluYrLeuSerLeuProIleAsnMet 180
DB 633 GCGTTTATGACTCTCTGGGGAATATAGCAGTTGACAGAGGTTGGGCTTACCAACAACA 692
QY 181 PheTrpValProTrpValTrpPheAlaAsnLeuSerMetYrAlaTrpLeuGlyValArg 200
DB 693 TTCGGGTACCCCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 752
QY 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGluMetAsnThrLeuArgTrpGln 220
DB 753 ATAGAGGACACGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 812
QY 221 CysGlyHisLeuYrAlaYrAspTrpIleSerIleProLeuValYrTrpGlyValVal 240
DB 813 TGTGACACCTATATGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 872
QY 241 ThrValAlaValYrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
DB 873 ACGATGGCGGTTTACAGCTTTTCTTCTGATGCTGATGCTGATGCTGATGCTGATGCT 932
QY 261 AlaValAlaYrProGlyHisGluLeuAspLeuValValProValPheThrPheLeuGln 280
DB 933 CTCAAGAACTACCCAGGCGCAGAGATGATGATGATGATGATGATGATGATGATGATG 992
QY 281 PhePhePheYrValGlyTrpLeuYrValAlaGluGlnLeuIleAsnProPheGlyGlu 300
DB 993 TTCATTATCTACATGCGCTGGCTCAAGGTGCGACAGATCAACACCTTTGGAGAG 1052
QY 301 AspAspAspAspPheGluThrAsnTrpIleValAspArgAsnLeuGlnValSerLeuLeu 320
DB 1053 GATGATGATGATTTTGAAGACCACTGATGATGATGATGATGATGATGATGATGATG 1112
QY 321 AlaValAspGluMetHisGlnAspLeuProArgMetGluProAspMetYrTrpAsnYr 340
DB 1113 TCTGGATGGAGATGACACAGAGCTGCTCCCTGAGGAGGATGATGATGATGATGATG 1172
QY 341 ProGluProGlnProProYrTrpTrpAlaIleSerAlaGlnPheArgArgAlaSerPheMet 360
DB 1173 GCGGCGGCTCAGCGGCTTACAGAGCTGCTTGGCGGGCTTTCACCGACCAATCTTCAAG 1232
QY 361 GlySerThrPheAsnIleSerLeuAsnYrGluGluMetGluPheGlnProAsnGlnGlu 380
DB 1233 GCGTCCACCTTCAACATGAGTCTGAAGAAAAAGACTCAGAGAGAGAGAGTCACTAT 1292
QY 381 Asp---GluGluAspAlaHisAlaGlyIleIleGlyArgPheLeuGlyLeuGlnSerHis 399
DB 1293 GACAAAGAAAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1352
QY 400 AspHisHisProProArgAlaAsnSerArgTrpYrYrLeuLeuTrpProYrAspArgGluSer 419
DB 1353 GACTACCATCTCTCCAAAGACAGACTTAAGACCAAACTGTGTGTCTTAAGAGAGACCC 1412

Qy	420	LeuHnH1SG1UG1yLeuPProLybasmH1b1ySa1a1a1yG1nAenVn1Arg1yG1n	433
Db	1413	TTCCCTC---GAGGACCACTCTTAAGAAATACAAGACGAAAGACCAAAA-----	1457
Qy	440	GlUaPaeuLySa1a1rPlLybLeuLySa1aVa1aPa1aPhelYseSerG1yProLeuTyr	459
Db	1458	-----AATGATGCTGTGGAAACCAAG-----GACCTCTTCAGAGCTATACCAAGTTT	1505
Qy	460	G1nArPProG1yTyrTyrSer1aPProG1nThrProLeuSerProThrPromePheDhe	479
Db	1506	AAGAAGCCAAATTACCGCCAGAGCCCCACAGACCACTGGCACGCCACTCCATGGTCTTC	1565
Qy	480	ProLeuG1nPProSer1aPProSer1yLybHnH1SeSerVal1ThnG1y1LaSPThrLySaP	499
Db	1566	CCTCCTCAGCAgTcAGCAcCCTCAgTcGCTTACTAGACACAGT-----	1610
Qy	500	LybSerLeuLyThrVaiSerSerG1y1aLybLybSerPheG1uLeuLeuSerG1uSer	519
Db	1610	-----	1610
Qy	520	AaPglYalALaUmeG1nH1aPProG1nVaiSerG1nVaiArgaGlyThrVaiG1nDhe	539
Db	1611	GGTGGGCTTCGCGACAGTACCAGAGGCGCTGTCACTACAGAAAGAAACAGTGGAGTTT	1670
Qy	540	AenLeuThrAaPmePProG1nU1aPProG1nasmH1SeLeuLybG1nProLeuG1nG1nSer	559
Db	1671	AACTTG---AACATTCCAGAGGCGCTTCAGAAACGTGAACCCCATTTGGACAG---	1724
Qy	560	ProThraen1LeH1aThrThrLeuLybAaP1aMeCaAProTyrTTPalALeUg1nAaSn	579
Db	1725	-----ATGCACACGATCTCTGAAGAGCGCTGGAGAACTCTATCCGTCAAGATATAA	1775
Qy	580	ArgaSP 581	
Db	1776	GCTGAC 1781	
RESULT 13			
BD136744			
LOCUS	BD136744	1916 bp	DNA linear PAT 18-SEP-2002
DEFINITION	Beet's macular dystrophy gene.		
ACCESSION	BD136744		
VERSION	BD136744.1	GI:23231689	
KEYWORDS	JP 2002504559-A/26.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Petrushkin, K., Caakey, T. C., Metzker, M. and Wadelius, C.		
TITLE	Beet's macular dystrophy gene		
JOURNAL	Patent: JP 2002504559-A 26 12-FEB-2002;		
COMMENT	MERCK & CO INC, CLAES WADLIUS		
	OS Mus musculus (mouse)		
	PN JP 2002504559-A/26		
	PD 12-FEB-2002		
	PF 25-FEB-1998 JP 2000533447		
	PR 60/075941,18-DEC-1998 US 62/112926 PT		
	KONSTANTIN PETRUSHKIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI		
	WADLIUS		
	PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P15/34, C12Q1/68//		
	PC C12P1/08,		
	PC C12N5/00, C12N15/00		
	CC Beet's macular dystrophy gene		
	EH key		
	FT 1..1916		
FEATURES	location/Qualifiers		
source	1..1916		
	location/Qualifiers		
	1..1916		
	/organism="Mus musculus (mouse)"		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10090"		
ORIGIN			

Alignment Scores:			
Prod. No.:	9,21e-159	Length:	1916
Score:	1899.00	Matches:	378
Percent Similarity:	74.32%	Conservative:	59
Best Local Similarity:	64.29%	Mismatches:	101
Query Match:	60.87%	Indels:	50
DB:	6	Gaps:	10
US-09-622-964A-3 (1-585) x BD136744 (1-1916)			
QY	1 MetThrIleThrTyrThrSerGlnValAlaAenAlaArgLeuGlySerPheSerArgLeu	20	
DB	11 ATGACTATACCTTACACCAACAAGTACCATGCCGCTGGTTCGTTCTGTCCTC	70	
QY	21 LeuLeuCySTPArgGlySerIleTyrIlySleuLeuTyrGlyGluPheLeuIlePheLeu	40	
DB	71 CTCCTGTGCTGGCGAGGAGCATCTTACAAGCTGCTGTATGAGAAATCTTGTCTTCAT	130	
QY	41 LeuCySTyrTyrIleIleArgPheIleTyrArgLeuAlaLeuThrGluGluGlnLeu	60	
DB	131 TTCTCTCTACTATTCATCCGTGGATCTTACAGAAATGGTCTCTCGATATCCGACGCTG	190	
QY	61 MetPheGlyIlySleuThrLeuTyrCySaaPseTyrIleGlnLeuIleProIleSerPhe	80	
DB	191 TTGTGTGAGAACTGAGCTCTGTACTGCGACACTTACATTACGCTCATCTCATCTTCTC	250	
QY	81 ValLeuGlyPheTyrValThrLeuValValThrArgTyrTPRAsnGlnTyrGluAenLeu	100	
DB	251 GTTCTGGGTTTCTATGTTACATTTGATGGAGACCCCTGCTGGAGCACATGAGAACTTTC	310	
QY	101 ProTyrProAspArgLeuMetSerLeuValSerGlyPheValGluGlyIlySaaPgluGln	120	
DB	311 CCGTGGCCCGACCCGCTCATGATCCAGAGTGTACGTTCTGTGAGGGCAAGATGAGGAA	370	
QY	121 SerArgLeuLeuArgArgThrLeuIleArgTyrAlaAenLeuGlyAaenValLeuIleLeu	140	
DB	371 GGCCCTTTGCTGGCGGCGACGCTCACCGCTACGCGACTCATCTGAGCCAAAGTCTCATCTCG	430	
QY	141 ArgSerValSerThrAlaValTyrIlySaaArgPheProSerAlaGlnIleSleuValGlnAla	160	
DB	431 CGGACATCATCAGACCTCGGTCTACAGACCTTTCCCACTCTTTCACCACTGTGTGCTAAGCA	490	
QY	161 GlyPheMetThrProAlaGlnIleIlySglnLeuGlnIlyLeuSerLeuProIleAaenMet	180	
DB	491 GGTTTATATGACCCATGGGGAACATTAAGCAGTTGAGAAATTGGGCTTACACACACACACA	550	
QY	181 PheTyrPValProTyrPValTyrPheAlaAenLeuSerMetIlyAlaTyrPheLeuGlyValArg	200	
DB	551 TTCTGGGTGGCCCTGGGTGTGTTTCCCAACTTGTCAATGAAGCCCTATTTTGGAGGTGCA	610	
QY	201 IleArgAspProIleLeuLeuGlnSerLeuLeuAaenGluMetAaenThrLeuArgTyrGln	220	
DB	611 ATCCGGGACACCGCTCTGCTCCAGAGCTCGATGAATGAGGTGTGACTTTTGGCTACTCAG	670	
QY	221 CysGlyIleSleuTyrAlaTyrAspTyrIleSerIleProLeuValTyrThrGlnValAla	240	
DB	671 TGTGACACAGCTGTATGCTTACAGACCTGGAATATATCCATTGCTGTATGACACAGCGGTGTG	730	
QY	241 ThrValAlaValTyrSerPhePheLeuThrCySleuValGlyArgGlnPheLeuAaenPro	260	
DB	731 ACAGTGGCAGATATACAGCTTTTCTTGATCTCTTGATCTGGAGGAGGAGTTTCTGAACCCA	790	
QY	261 AlaIlySaaIleTyrProGlyIleGluLeuAaenPheLeuValProValPheThrPheLeuGln	280	
DB	791 AACAAAGACTACCCAGGCGCATGAGATGAGATCTGGGTGTGCTGTCTTCAAAATCTTCGAA	850	
QY	281 PhePhePheTyrValGlyTyrPheIlySaaIleGluGlnLeuIleAaenProPheGlyGlu	300	
DB	851 TTCTTATTTCTATCATATGGGCTGGCTGAAGGTGGCAAGCAAGCTCATCAACCCCTTGGGGAG	910	
QY	301 AspAspAspAspPheGluThrAsnTyrIleValAspArgAaenLeuGlnValSerLeuLeu	320	
DB	911 GACGATGTATGATTTTGAAGACTAACTGATCATATGTACAGAAACCTGAGGTGTCCCTGTGTG	970	

```

Qy 321 AlavAlaAspGluMetHisGlnAspLeuProArgMetGluProAspMetTyrTrpAsnLys 340
Db 971 TCCTGGATGGGATGTCACCAAGATGCTCTCCAGACGTCATGATCTGGAACGAG 1030
Qy 341 ProGluProGlnProProTyrTrpAlaAlaSerAlaGlnPheArgAlaSerPheMet 360
Db 1031 GCACGGCTCAGCGCCCTACAGCGCTTCTCCAGAGCTCGCCGACATTCCTTCAG 1090
Qy 361 GlysSerThrPheAsnIleSerLeuAsnLysGluGluMetGluPheGlnProAsnGlnLys 380
Db 1091 GGCTCCACCTTCACATCAGCTTAAGAGAGAGACTTAAGAGCTTTGCTCAAAAGAGAG 1150
Qy 381 -----AspGluGluAspAlaHisAlaGlyIleIleGlyLysPheLeuGlyLeu 396
Db 1151 GCTGACACGAGTAAAGAAAGAGAGCTATAGACAGACCTAGAGCTGCTTCTAGGACTG 1210
Qy 397 GlnSerHisAspHisSerProProArgAlaAsnSerArgThrLysLeuLeuTrpProLys 416
Db 1211 CAACCCMAAACTACCATCTTCTCCGMAAGACTTAAGACCAACTATTTGTTCTTAG 1270
Qy 417 ArgGluSerLeuLeuHisGluGlyLeuProLysAsnHisLysAlaAlaLysGlnSerVal 436
Db 1271 AACCCCTCTC-----GAGGCCAGCTTAAGGAT-----GCCAACGAG 1309
Qy 437 ArgGlyGlnGluAspAsnLysAlaTrpLysLeuLysAlaValAspAlaPheLysSerGly 456
Db 1310 AAAAACACAGAAAGAT-----GCTGGAATTTAAGGCTGAGCTTCTGAAATGCTT 1363
Qy 457 ProLeuTrpGlnArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrPro 476
Db 1364 CCAAGTTTAAGAGAGAGGCTCCATCTGTCGCCACAGCCAGCCAGCAGC----- 1414
Qy 477 MetPhePheProLeuGluProSerAlaProSerLysLeuHisSerValThrGlyIleAsp 496
Db 1415 -----CACCTTACTGAGCAGCTCAGACCTCTCAGT-----TCAGACACAGCT 1456
Qy 497 ThrLysAspLysSerLeuLysThrValSerSerGlyAlaLysLysSerPheGluLeu 516
Db 1456 ----- 1456
Qy 517 SerGluSerAspGlyAlaLeuMetGluHisProGluValSerGlnValArgArgLysThr 536
Db 1457 -----GATGGCCCTTCCACAGATTAACAAGAAATCTGTCACTCAAAAAAAACT 1507
Qy 537 ValGluPheAsnLeuThrAspMetProGluIleProGluAsnHisLysLysGlu---Pro 555
Db 1508 GTGAGTTTAACCTG---AACATTCAGAGAGAGCCCAAGAAACATCTTCAACAGCGCGT 1564
Qy 556 LeuGluGlnSerProThrAsnIleHisThrThrLeuLysAspHisMetAspProTyrTrp 575
Db 1565 TTGACCCAGATGTCACCAATATACAGGCTTAAATGAAGAGCATGACAGATCCTAT--- 1621
Qy 576 AlaLeuGluAsnArgAspGluAla 583
Db 1622 -----CCCTACAGGATGAAGCT 1639

```

```

RESULT 14
LOCUS AY450427 1904 bp mRNA linear ROD 24-JUN-2004
DEFINITION Mus musculus vitelliform macular dystrophy 2 (Vmd2) mRNA, complete
cDNA.
ACCESSION AY450427
VERSION AY450427.1 GI:41612986
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Stohr, H., Marguard, A., Nanda, I., Schmidt, M. and Weber, B.H.
Three novel human VMD2-like genes are members of the evolutionary
highly conserved RFP-TM family

```

```

JOURNAL Eur. J. Hum. Genet. 10 (4), 281-284 (2002)
MEDLINE 22027749
PUBMED 12032738
REFERENCE 2 (bases 1 to 1904)
AUTHORS Kraemer, F., Stoehr, H. and Weber, B.H.F.
JOURNAL Cloning and characterization of the murine Vmd2 RFP-TM gene family
REFERENCE Cytogenet. Genome Res. 105 (1), 107-114 (2004)
AUTHORS 3 (bases 1 to 1904)
TITLES Weber, B.H.F. and Kraemer, F.
JOURNAL Submitted (27-OCT-2003) Human Genetics, University of Wuerzburg, Am
Hubland, Biozentrum, Wuerzburg 97074, Germany
FEATURES
source
1..1904
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="19"
/tissue="testis"
1..1904
/gene="Vmd2"
37..1692
/gene="Vmd2"
/codon_start=1
/product="vitelliform macular dystrophy 2"
/protein_id="AA09922.1"
/db_xref="GI:41612989"
/translacion="MTITVTKVAVNARIGSFSSLLCWRGSIYKLLGEPVFTLY
SIRGLRYMTLSDPOLPEKLLAYCDSTYLPIPIFVAGFYTTLVSRMSQENLPM
PRLMTQVSSPVEGDEBGRRLRRLIRAILGVLIRISITSVYKRPPTLHLVLA
GPWTHREHQLQKLGPHNTEVWVWVWFENLSMKYILGRLRDVLQSLNUEVCTLR
TQGGQILADWISILVVTQVTVAVSEFLACLIGQFLPNPKDYPGHEMDLVVPR
TLIQFLYMGMLKVAOLINPGEHDDDEFTWILIDRMLOVLSLSDGMHONLPMER
DYMWEAPQPYTAASARSRRHSFMGSTFNLSEKEDLELMSKEADTDKRGYS
TTCGFLGLOPKNYHLPLKDLTKKLCSKNPLLEGCKRANOKMDKWKPKGLDPLK
VPRFKRGHGCPQAPSSHPTEOSA PSSSDPGDGSSTYQETCHKKKTVERPNLIP
SPTHLQQRRLQMSNTNQLMKHAESEYPRDEAGTPEVLE"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 1,69e-158 Length: 1904
Score: 1896.00 Matches: 377
Percent Similarity: 74.32% Conservative: 60
Best Local Similarity: 64.12% Mismatches: 101
Query Match: 60.77% Indels: 50
DB: 10 Gaps: 10

```

US-09-622-964A-3 (1-585) x AY450427 (1-1904)

```

Qy 1 MetThrIleThrTyrThrSerGluValAlaAsnAlaArgLeuGlySerPheSerArgLeu 20
Db 37 ATGACTATACCTTACACAAACAAGATGACCAATCCCGCTCGGTTGCTCGTCCCTC 96
Qy 21 LeuLeuGlyTyrPArgLysSerLysLysLeuLeuTyrGlyGluPheLeuIlePheLeu 40
Db 97 CTCCTGCTGCGCAGGAGCAGCATCAACAGCTGCTGATGAGAAATTCCTTGTCTATA 156
Qy 41 LeuGlyTyrTrpIleIleArgPheIleTyrArgLeuAlaLeuThrGluGlnGlnLeu 60
Db 157 TTCCTCTACTATTCATCCATCCGTCGACTCTACAGATGTTCTCTCAGTGTACGAGCTG 216
Qy 61 MetPheGluLysLeuThrLeuTyrCysAspSerTyrIleGlnLeuIlePheIleSerPhe 80
Db 217 TTGTTTGAAGAGCTGCGCTGTACTGCGACAGCTACATCCAGCTCATCCCTATATCTTC 276
Qy 81 ValLeuGlyPheTyrValThrLeuValValThrArgTrpPheAsnGlnTyrGluAsnLeu 100
Db 277 GTTCTGGATTCTATATGTTACATTTGTTGAGCGCGCTGGTGGAGCCAGTACGAACTTG 336
Qy 101 ProTrpProAspArgLeuMetSerLeuValSerGlyPheValGluGlyLysAspGluGln 120
Db 337 CCGTGGCCCAACGCTCATGATCCAGGTGTCTAGCTTGTGTGAGGCAAGATGAGAA 396

```

```

Qy 121 SerArgLeuLeuAArgThrLeuIlleArgTyrAlaIleuLeuGlyAsnValLeuIlleu 140
Db 397 GCGCGTTTGTCGCGCGGACCGCTCATCCGCTACCGCATCTCGGCGCAAGCTCTATCTTG 456
Qy 141 ArgSerValSerThrAlaValTyrTlyValArgPheProSerAlaGlnHisLeuValGlnAla 160
Db 457 CGCAGCATCAGACCTCGGTCTACAAAGCGCTTCCCATCTTCCACCACTCGTGTGTAGCA 516
Qy 161 GlyPheMetThrProAlaGlnHisGlyValGlnLeuGlyLeuSerLeuProHisAsnMet 180
Db 517 GGTTTTAGACCCCATGGGGAACATAGCATGGAAGTGGGCGCTACCAACACACCA 576
Qy 181 PheTyrValProTyrValTyrPheAlaAsnLeuSerMetLeuValTyrTyrGlyValArg 200
Db 577 TTCGCGGCGCTCGGTGTGTTTCCCACTTGCTCAATGAAGCGCTATCTTGAGGTGCA 636
Qy 201 IleArgAspProIleLeuLeuGlnSerLeuLeuAsnGlnMetAsnThrLeuArgThrGln 220
Db 637 ATCCGGGACACCGTCTGCTCCAGAGCTGATGATGATGATGATGATGATGATGATGATGAT 696
Qy 221 CysGlyHisLeuTyrAlaTyrAspTyrIleSerIleProLeuValTyrThrGlnValVal 240
Db 697 TGTGACAGCTGTATGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 756
Qy 241 ThrValAlaValTyrSerPhePheLeuThrCysLeuValGlyArgGlnPheLeuAsnPro 260
Db 757 ACAGTGGCAGTATACAGCTTTTCTTGATGATGATGATGATGATGATGATGATGATGATGATG 816
Qy 261 AlaValAlaValTyrProGlnHisGlnLeuAspLeuValAlaProValPheThrPheLeuGln 280
Db 817 AACAGGACATACCCAGGCGCATGAGATGATGATGATGATGATGATGATGATGATGATGATG 876
Qy 281 PhePhePheTyrValGlyTyrLeuValAlaGlnGlnLeuIleAsnProPheGlyGlu 300
Db 877 TTCCTTATCTTACATGAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 936
Qy 301 AspAspAspAspPheGlnThrAsnTyrIleValAlaAspAsnLeuGlnValSerLeuLeu 320
Db 937 GACCATGATGATTTTGACATCTTACATGATGATGATGATGATGATGATGATGATGATGATG 996
Qy 321 AlaValAlaValGlnMetHisGlnAspLeuProArgMetGluProAspMetTyrTyrAsnHis 340
Db 997 TCCGTGATGGGATGCAACGACCTGCTCCCATGGAAGCTGACATGATGATGATGATGATG 1056
Qy 341 ProGlnProGlnProProTyrThrAlaAlaSerAlaGlnPheArgArgAlaSerPheMet 360
Db 1057 GCAAGCGCTCAGCGCGCTTACAGAGCTGCTTCCAGAGTCTCGCGCATTCCTTCATG 1116
Qy 361 GlySerThrPheAsnIleSerLeuAsnHisGlnGlnMetGlnPheGlnProAsnGlnGlu 380
Db 1117 GCGTCCACCTTCAACATCAGCTTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1176
Qy 381 -----AerGlnGlnAspAlaHisIleAlaGlyIleIleGlyArgPheLeuGlyLeu 396
Db 1177 GCTGACACGATAGAAAGAGAGTGGCTATAGCAGACACCAAGCTGCTTCTTAGAGCTG 1236
Qy 397 GlnSerHisAspHisHisProProArgAlaAsnSerArgThrTyrLeuLeuTyrProLys 416
Db 1237 CAACCCAAAACATCCTCTCCCTTGAAAGACTTAAAGACCAACATCTGTGTTCTTAAG 1296
Qy 417 ArgGlnSerLeuLeuHisGlnGlyLeuProLysAsnHisGlyValAlaValGlnAsnVal 436
Db 1297 AACCCCTCTCTC-----GAAGGCCAGCTGAAGAT-----GCCAACAG 1335
Qy 437 ArgGlyGlnGlnAspAsnLysAlaTyrPheLeuValValAlaAspAlaPheLysSerGly 456
Db 1336 AAAAACCAAGAAAGT-----GTCTGAAATTTAAAGGTCTGCACTTCTGAATGTGCTT 1389
Qy 457 ProLeuTyrGlnArgProGlyTyrTyrSerAlaProGlnThrProLeuSerProThrPro 476
Db 1390 CCAAGGTTTAAAGAGAGAGGCTCCATGTGTGCGCCACAGGAGCAGCAGCAGC----- 1440

```

```

Qy 477 MetPhePheProLeuGlnProSerAlaProSerTyrLeuHisSerValThrGlyIleAsp 496
Db 1441 -----CACCTTACTGAGCAGTACAGACCCCTCCAGT-----TCAGACACAGGT----- 1482
Qy 497 ThrLysAspLysSerLeuLysThrValSerSerGlyAlaLysLysSerPheGlnLeuLeu 516
Db 1482 ----- 1482
Qy 517 SerGlnSerArgGlyAlaLeuMetGlnHisProGlnValAlaSerGlnValArgArgLysPhe 536
Db 1483 -----GATGGCGCTTCCACAGATTAACCAAGAAATCTGTGCATGATAAAGAAAGAACT 1533
Qy 537 ValGlnPheAsnLeuThrAspMetProGlnIleProGlnAsnHisLeuLysGlu---Pro 555
Db 1534 GTGAGATTAACTTG---AACATTCAGAGAGCCCCACAGAAACATCTTCAACAGCGCGCT 1590
Qy 556 LeuGlnLysSerProThrAsnIleHisThrThrLeuLysAspHisMetAspProTyrTyr 575
Db 1591 TTGGACCAAGATGTCAACCAATATGACAGCTCTAATGAAAGAGCATGACAGACTCTAT--- 1647
Qy 576 AlaLeuGlnAsnArgAspGlyAla 583
Db 1648 -----CCTTACAGGATGAAGCT 1665

RESULT 15
LOCUS AY064707 1289 bp mRNA linear MAM 19-AUG-2002
DEFINITION Sus scrofa beestrophin mRNA, partial cds.
ACCESSION AY064707
VERSION AY064707.1 GI:18476495
KEYWORDS
SOURCE
ORGANISM
Sus scrofa (pig)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
1 (bases 1 to 1289)
Marmorstein, L.Y., McLaughlin, P.J., Stanton, J.B., Yan, L., Crabb, J.W.
and Marmorstein, A.D.
Beestrophin interacts physically and functionally with Protein
Phosphatase 2A
J. Biol. Chem. 277 (34), 30591-30597 (2002)
JOURNAL
J. Biol. 12058047
PUBMED
2 (bases 1 to 1289)
Marmorstein, L.Y., McLaughlin, P.J., Stanton, B., Yan, L., Crabb, J.W.
and Marmorstein, A.D.
Direct Submision
Submitted (30-NOV-2001) Cole Eye Institute, 331, Cleveland Clinic
Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA
FEATURES
source
1..1289
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
<1..1289
/codon_start=3
/product="beestrophin"
/protein_id="AAU40882.1"
/db_xref="GI:18476496"
/translation="PQHLYVAGFMTPESEHNLKELSLPHNSFMKPMWVPMNLSTKANI
GGRILRPVLYOSLIDENMTLTTCGHLUADMSVPLVYQVTVAVSFFLACTVVR
OFLNPKAVYGNHMDLVVPLFTLOFFYVGMVKVABOLINPGEEDDDPEFTMIYVR
SLQVLSAIDEMHNDLRPMERDMYNDREHPRYTASASRRSPFGSGTPTSLGKE
DMERQFVEEENATGILGNLGLQSSDHPRPRTNSKTKLMPKKGHFNHGNPLNG
ARLDDSSQEDSKMREGGFSAALCSRGVNSAPQVTLGHTPMVFPPEESAPGLKRV
SGIDEAKQSLQPATPISIKKSFELPESAEASAPLOGSHVIRKTVIEFNILADSEAP
EHLKEPILBPPMGIHAIILKHDRDPYALLENRDBANS"
ORIGIN
Alignment Scores:
Pred. No.: 2..2e-141 Length: 1289
Score: 1701.00 Matches: 324
Percent Similarity: 81.52% Conservative: 29
Best Local Similarity: 74.83% Mismatches: 72

```

Query Match: 54.52% Indels: 8
 DB: 4 Gaps: 4
 US-09-622-964A-3 (1-585) x AY064707 (1-1289)

```

QY 155 GlnHisLeuValGlnAlaGlyPheMetThrProAlaGlnHisLeuGlnLeuGlnLeu 174
DB 6 CAACACCTGGTGAAGAAGAGGCTTATAGACCCCTCGGAAACACAGACCTTAGAAAACTG 65
QY 175 SerLeuProHisAenMetPheTrpValProTrpValTrpPheAlaSerLeuSerMetLys 194
DB 66 AGCTTGCACACCAACTATCTGGATGCCCTGGGTGGTGGTTCACAACTGTTCACAGAAAG 125
QY 195 AlaTrpLeuGlnGlyArgLysLeuArgAspProIleLeuLeuGlnSerLeuLeuGlnMet 214
DB 126 GCCGTGATGGGGGTCAATCCGGGACCCGTGCTCTCCAGAGTGTGCTTCACAGAAATG 185
QY 215 AsnThrLeuArgThrGlnGlyArgLysLeuValAlaTrpAspTrpLysSerLeuProLeu 234
DB 186 AACACCTTGGCTTACGTGTGAGACACCTGTATGCTTACGACCTGATCGATGTCCTCGCTG 245
QY 235 ValTrpThrGlnValValThrValAlaValTyrSerPhePheLeuThrCysLeuValGly 254
DB 246 GTGTACACTCAGGTGTGACCGGTGGCGGTATACACTTCTTCTGCTGTGTGTGAG 305
QY 255 ArgGlnPheLeuAsnProAlaLysAlaTyrProGlnHisGlnLeuAspLeuValPro 274
DB 306 AGGCAATTTCTGAACCCACCAAGGCTTACCCCGGCATGATGATGACCTGTGTACCC 365
QY 275 ValPheThrPheLeuGlnPhePhePheTyrValGlyTrpLeuValAlaGlnLeu 294
DB 366 CTCTTCAAGTCTCTGACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 425
QY 295 IleAsnProPheGlnGlyLeuAspAspPheGlnThrAsnTrpIleValAspArgAsn 314
DB 426 ATCAACCCCTTGGAGAGATGATGACGACTTGGAGACCACTGSAATTGTGACAGAGAGC 485
QY 315 LeuGlnValSerLeuLeuAlaValAspGlnMetHisGlnAspLeuProArgMetGluPro 334
DB 486 TTGCAGGTGTCTCTTCTGCTGTATGATGATGACACGACCTGCGCACCCATGAGAGCA 545
QY 335 AspMetTyrTrpAsnLysProGlnProGlnProTyrTrpAlaLysSerAlaGlnPhe 354
DB 546 GATATGTACTGGAATGATCCAGAACCAATCCCTTACACAGCTGCTTCTGCCAGTCT 605
QY 355 ArgArgAlaSerPheMetGlySerThrPheAsnIleSerLeuAsnLysGlnGlnMetGln 374
DB 606 CGCCGACCTCTCTTCTTCTGCTTCCACTTCAATATCAGTCTGGGTAAAGAGACATGGAG 665
QY 375 PheGlnProAsnGlnGlnAspGlnGlnAspAlaHisAlaGlyIleIleGlyArgPheLeu 394
DB 666 TTCCAGCGG-----GAGGAGAGAGAGAGAGCGCACCGGCACTCTGGCACTTCTCTG 719
QY 395 GlyLeuGlnSerHisAspHisHisProProArgAlaAsnSerArgThrLysLeuLeuTrp 414
DB 720 GGGCTGCAATCCACGACCAACAGCCCTCCAGCAAACTCAAGAGCAAACTCTCTGTGG 779
QY 415 ProLysArgGlnSerLeuLeuHisGlnGlyLeuProLysAsnHisLysValAlaLysGln 434
DB 780 CCCAAGAAAGAGGCCCTTTCACAGAGGGCCACCCCAAGAACTCAAGGGGGGCCAGGCTG 839
QY 435 AsnValArgGlnGlnLysAspAsnLysAlaTrpLysLeuLysValAlaAspAlaPheLys 454
DB 840 GACTCTAGCAGACGAAAGACAGAGGCTTGAGG-----GAGGTGCTTCAAG 890
QY 455 SerGlyProLeuTrpGlnArgProGlyTyrTyrSerAlaProGlnThrProLeuSerPro 474
DB 891 TCCGCTGCTCTGTGTGGAGGCGGGGTACACAGCGGCCCAAGACACCTCTCGGCCAC 950
QY 475 ThrProMetPhePheProLeuGlnProSerAlaProSerLysLeuHisSerValThrGly 494
DB 951 ACCCTATGGTCTTCCACCCGAAAGAGTCAAGCCCTTAGGTCTTGGCAGAGTCTGGGAC 1010

```

```

QY 495 IleAsp-----ThrLysAspLysSerLeuLysThrValSerSerGlyAlaLysLysSer 512
DB 1011 ATAGATGAGGCTGCAAGAACAGAGGCTTACGCTTCCAGCTCCAGACATCAAGAAAGAT 1070
QY 513 PheGlnLeuLeuSerGlnSerAspGlyAlaLeuMetGlnHisProGlnValSerGlnVal 532
DB 1071 TTCAGTGTCTCCAGAGAGGCGCGAGGCTCGGCGGAGCCCTCGCAAGGAGTCAAGT 1130
QY 533 ArgArgLysThrValGlnPheAsnLeuThrAspMetProGlnLysProGlnAsnHisLeu 552
DB 1131 AGAGAGAAAGAGCGTGAAGTTTAACTGGCCGATCTGTGAGAGGCCCGAA--CATCTC 1187
QY 553 LysGlnProLeuGlnGlnSerProThrAsnIleHisThrThrLeuLysAspHisMetAsp 572
DB 1188 AAAAGAACGAATTTGGAGCCACCGATGGGCATGACGATATCTCAAGACCAAGCGGAT 1247
QY 573 ProTyrTrpAlaLeuGlnAsnArgAspGlnAlaHisSer 585
DB 1248 CCTACTGGGCTTGGAAACAGGAGATGAAGACACTTCC 1286

```

Search completed: March 23, 2005, 15:22:03
 Job time : 7505.71 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 22:00:09 ; Search time 10345.6 Seconds
(without alignments)
11376.592 Million cell updates/sec

Title: US-09-622-964a-4
Perfect score: 2429
Sequence: 1 CAGGAGATCCACCCAGCCTA.....AAAAAAAAAAAAAAAAAAAA 2429

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_to:*
11: gb_str:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2429	100.0	2429	6	BD136721
2	2418.4	99.6	2420	6	AF057170 Homo sapi
3	2013	82.9	2229	6	BD136720 Beest's ma
4	1995.4	82.1	2210	9	AF057169 Homo sapi
5	1953.2	80.4	2171	9	AF073501 Homo sapi
6	1942.8	80.0	2170	6	CQ729633 Sequence
7	1809.6	74.5	2187	9	AY357925 Macaca fa
8	1587	65.3	2441	9	BC015220 Homo sapi
9	1543.4	63.5	1758	9	AY515704 Homo sapi
10	1541.8	63.5	1758	6	AX745964 Sequence
11	1371	56.4	2435	6	BC041664 Homo sapi
12	1244.2	51.2	1263	6	BD132659 Secreted
13	712.8	29.3	2035	10	BC079048 Ratcun no
14	680.4	28.0	1904	10	AY450427 Mus muscu
15	658.2	27.1	1916	6	BD136744 Beest's ma
16	648.4	26.7	1289	4	AY064707 Sub scrof
17	642.2	26.4	16125	6	BD136719 Beest's ma
18	642.2	26.4	142092	9	AF139813 Homo sapi
19	642.2	26.4	163024	9	AP006260 Homo sapi

C	20	642.2	26.4	163915	2	AC087451	AC087451 Homo sapi
	21	642.2	26.4	166867	9	AP003733	AP003733 Homo sapi
	22	642.2	26.4	196080	9	AC004228	AC004228 Homo sapi
	23	640.6	26.4	706	9	HSVMD2P10	AF073499 Homo sapi
	24	640.6	26.4	112309	9	AC003025	AC003025 Human chr
	25	640.6	26.4	133683	2	AC084857	AC084857 Homo sapi
	26	599	24.7	726	9	AY357926810	AY357935 Macaca fa
C	27	552.8	22.8	160169	2	AC051664	AC051664 Homo sapi
	28	431.4	17.8	2137	6	AX745966	AX745966 Sequence
	29	431	17.7	1530	6	C0731444	C0731444 Sequence
	30	431	17.7	1530	6	AX745993	AX745993 Sequence
	31	431	17.7	1530	9	AY515705	AY515705 Homo sapi
	32	431	17.7	1908	9	AF440756	AF440756 Homo sapi
	33	415	17.1	1954	10	AY450428	AY450428 Mus muscu
	34	415	17.1	1956	10	BC031186	BC031186 Mus muscu
	35	415	17.1	1957	10	BC019528	BC019528 Mus muscu
	36	415	17.1	1992	10	BC036157	BC036157 Mus muscu
	37	389.2	16.0	1861	9	AK000139	AK000139 Homo sapi
	38	369.4	15.2	1898	5	BC043854	BC043854 Xenopus l
	39	369.4	15.2	1921	5	AY273825	AY273825 Xenopus l
	40	365.2	15.0	1882	5	CR760914	CR760914 Xenopus t
	41	365.2	15.0	2021	5	BC061379	BC061379 Xenopus t
	42	365.2	15.0	2021	5	HSVMD2P11	AF073500 Homo sapi
C	43	358.4	14.8	492	9	AC004588	AC004588 Homo sapi
	44	358.4	14.8	106648	2	BC0844229	BC0844229 Xenopus l
	45	358.2	14.7	2064	5	BC0844229	BC0844229 Xenopus l

ALIGNMENTS

RESULT 1
LOCUS BD136721 2429 bp DNA linear PAT 18-SEP-2002
DEFINITION Beest's macular dystrophy gene.
BD136721
ACCESSION BD136721.1 GI:23231666
VERSION JP 2002504559-A/3.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Petrunkhin,K., Caskey,T.C., Metzker,M. and Wadelius,C.
1 (bases 1 to 2429)
REFERENCE
Petrunkhin,K., Caskey,T.C., Metzker,M. and Wadelius,C.
Patent: JP 2002504559-A 3 12-FEB-2002;
JOURNAL MERCK & CO INC, CLAES WADLIUS
OS Homo sapiens (human)
PN JP 2002504559-A/3
PD 12-FEB-2002
PR 22-FEB-1999 JP 2000533447
PR 25-FEB-1998 US 60/075941,18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUNKHIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI
WADLIUS

PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68 // PC
C12P21/08,
PC C12N5/00, C12N15/00
CC Beest's macular dystrophy gene
FH key location/Qualifiers
FT source 1..2429
FT location/Qualifiers
FT /organism='Homo sapiens (human)'.
FT /organism='Homo sapiens'
FT /mol_type='genomic DNA'
FT /db_xref='taxon:9606'

FEATURES

source

ORIGIN

Query Match 100.0%; Score 2429; DB 6; Length 2429;
Beest Local Similarity 100.0%; Pred. No. 0;
Matches 2429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CAGGAGATCCACCCAGCCTTGTGGATCATCGACCCACCTGGAA 60
Oy


```
Db      1 CAGGAGTCCACAGCTAGTCCGACACTTTGTGGGATCATGTGGACCCACCTGGAA 60
Qy      61 CCCCACTGACCCAGCCCACTGTGCAAGCCCACTGCGCATGACCATCACTTA 120
Db      61 CCCCACTGACCCAGCCCACTGTGCAAGCCCACTGCGCATGACCATCACTTA 120
Qy      121 CAAGCCAAGTGGCTAATAGCCGCTTAGGCTCTTCCGCCGCTGCTGTGTGTCGGG 180
Db      121 CAAGCCAAGTGGCTAATAGCCGCTTAGGCTCTTCCGCCGCTGCTGTGTGTCGGG 180
Qy      181 GCAGCATCTACAAAGCTCTATATGGGAGTTCTTAATCTTCGCTCTGCTACTATCA 240
Db      181 GCAGCATCTACAAAGCTCTATATGGGAGTTCTTAATCTTCGCTCTGCTACTATCA 240
Qy      241 TCGGCTTTATTTATAGGCTGGCCCTCAAGGAAACAAGCTGATGTTTGAAAATGA 300
Db      241 TCGGCTTTATTTATAGGCTGGCCCTCAAGGAAACAAGCTGATGTTTGAAAATGA 300
Qy      301 CTCTGATTGCGACAGCTACATCAGCTCATCCCAATTCCTTCGCTGGGCTTCTAC 360
Db      301 CTCTGATTGCGACAGCTACATCAGCTCATCCCAATTCCTTCGCTGGGCTTCTAC 360
Qy      361 TGAAGCTGTGTGACCCCGTGTGGACCAAGTACGAAACCTGCGTGGCCGACCGC 420
Db      361 TGAAGCTGTGTGACCCCGTGTGGACCAAGTACGAAACCTGCGTGGCCGACCGC 420
Qy      421 TCATGAGCTGGTGGGGCTTCGTGAAGGCAAGGACGAGAACCGCGCTGTGCGGC 480
Db      421 TCATGAGCTGGTGGGGCTTCGTGAAGGCAAGGACGAGAACCGCGCTGTGCGGC 480
Qy      481 GCAAGCTCATCCGCTACAGCCAACTTGGGCAAGTCTCATCTTGGCAGAGCTCAG 540
Db      481 GCAAGCTCATCCGCTACAGCCAACTTGGGCAAGTCTCATCTTGGCAGAGCTCAG 540
Qy      541 CAGTCAACAAGGGCTTCCCAAGCGCCAGCAACTGTGTGCAAGCAGGCTTTATGA 600
Db      541 CAGTCAACAAGGGCTTCCCAAGCGCCAGCAACTGTGTGCAAGCAGGCTTTATGA 600
Qy      601 CAGAACACAAGCAATGGAGAACTGAGCTTACACACAAATGTTCTGGGTGGCC 660
Db      601 CAGAACACAAGCAATGGAGAACTGAGCTTACACACAAATGTTCTGGGTGGCC 660
Qy      661 TGTGCTTGGCCAACTGTCATGAAGGCTGTGGAGTGAATCCGGGACCTTATCC 720
Db      661 TGTGCTTGGCCAACTGTCATGAAGGCTGTGGAGTGAATCCGGGACCTTATCC 720
Qy      721 TGCTCAGAGCTGTGCTGAACAGATGAACCTTGCGTACATGATGAGACCTGTAT 780
Db      721 TGCTCAGAGCTGTGCTGAACAGATGAACCTTGCGTACATGATGAGACCTGTAT 780
Qy      781 CCTACGACTGATTAATATCCACTGTGTATACACAGTGTGACTGTGGCGGTGTA 840
Db      781 CCTACGACTGATTAATATCCACTGTGTATACACAGTGTGACTGTGGCGGTGTA 840
Qy      841 GCTTCTTCTGATCTGTCTAGTGGGGGGGAGTTTCTGAACCCAGCCAAAGGCTTA 900
Db      841 GCTTCTTCTGATCTGTCTAGTGGGGGGGAGTTTCTGAACCCAGCCAAAGGCTTA 900
Qy      901 GCATGAGCTGAGCCTGTTGTGCCGTCTTCAAGTTCACTGAGTTCTTCTATATG 960
Db      901 GCATGAGCTGAGCCTGTTGTGCCGTCTTCAAGTTCACTGAGTTCTTCTATATG 960
Qy      961 GCTGAGCTGAAGTGGGCTCTCCAGGACCTGCTGGGCTGAGGCAATGGCAGAGG 1020
Db      961 GCTGAGCTGAAGTGGGCTCTCCAGGACCTGCTGGGCTGAGGCAATGGCAGAGG 1020
Qy      1021 ATGGCCAGAGCTGCTTGAAGCAGAGATGAGTGTCAAGAAAGTCTCAGGGTATG 1080
Db      1021 ATGGCCAGAGCTGCTTGAAGCAGAGATGAGTGTCAAGAAAGTCTCAGGGTATG 1080
Qy      1081 AAAGCAGCCAGGCGTGTGGCACAACCTGTAAATCCAGCTACTCGGAGAGCTGAG 1140
Db      1081 AAAGCAGCCAGGCGTGTGGCACAACCTGTAAATCCAGCTACTCGGAGAGCTGAG 1140
Qy      1141 GAGAAATGCTTGAACCCGGGGAGGGAGTTGTGTGGCAGAGAGCTCATCAACCCCT 1200
Db      1141 GAGAAATGCTTGAACCCGGGGAGGGAGTTGTGTGGCAGAGAGCTCATCAACCCCT 1200
Qy      1201 TGAAGAGATGATGATATTTTGAAGCAAACTGATTTGTGACAGAAATTTGAGGT 1260
Db      1201 TGAAGAGATGATGATATTTTGAAGCAAACTGATTTGTGACAGAAATTTGAGGT 1260
Qy      1261 CCTGTGGCTGTGATGATGATGACAGAGACCTGCTCGGATGAGCCGACATGTACT 1320
Db      1261 CCTGTGGCTGTGATGATGATGACAGAGACCTGCTCGGATGAGCCGACATGTACT 1320
Qy      1321 GAATTAAGCCGAGCAGAGCCCTTACACAGGCTTCCGCCAGTCCGTGTGAGACCT 1380
Db      1321 GAATTAAGCCGAGCAGAGCCCTTACACAGGCTTCCGCCAGTCCGTGTGAGACCT 1380
Qy      1381 CTTTATGGGCTCCACTTCAACATCAGCTTGAACAAAGAGAGATGAGTTCCAGCC 1440
Db      1381 CTTTATGGGCTCCACTTCAACATCAGCTTGAACAAAGAGAGATGAGTTCCAGCC 1440
Qy      1441 TCAAGAGAGAGAGAGATGCTCAGCTGGCATATTTGGCCGCTTCTAGGCTGTGAT 1500
Db      1441 TCAAGAGAGAGAGAGATGCTCAGCTGGCATATTTGGCCGCTTCTAGGCTGTGAT 1500
Qy      1501 CCATGATACCATCTTCCAGAGGCAAACTCAAGACCAAACTTACTGTGSCCAAGAG 1560
Db      1501 CCATGATACCATCTTCCAGAGGCAAACTCAAGACCAAACTTACTGTGSCCAAGAG 1560
Qy      1561 ATCCCTTCTCCAGAGAGGCTGCCCAAAAACAAGAGAGGCAAGAGCTTATGAG 1620
Db      1561 ATCCCTTCTCCAGAGAGGCTGCCCAAAAACAAGAGAGGCAAGAGCTTATGAG 1620
Qy      1621 CCAGAGAGACAACAAGGCTTGAAGGCTGTGAGCGCTTCAAGTCTGAGCCACT 1680
Db      1621 CCAGAGAGACAACAAGGCTTGAAGGCTGTGAGCGCTTCAAGTCTGAGCCACT 1680
Qy      1681 GTATCAGAGGCGGCTTACATCAGTGCACACAGAGCGCCCTCAGCCCACTCCA 1740
Db      1681 GTATCAGAGGCGGCTTACATCAGTGCACACAGAGCGCCCTCAGCCCACTCCA 1740
Qy      1741 CTTCCCTTGAAGAACATCAGGCGCGTCAAAAGTTTCAAGTCTCAGGCAATGAG 1800
Db      1741 CTTCCCTTGAAGAACATCAGGCGCGTCAAAAGTTTCAAGTCTCAGGCAATGAG 1800
Qy      1801 AAGCAAAAGCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATGCT 1860
Db      1801 AAGCAAAAGCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATGCT 1860
Qy      1861 GAGCGATGGGGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAAAT 1920
Db      1861 GAGCGATGGGGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAAAT 1920
Qy      1921 GTTTAACTGACGATATGACAGAGATCCCGAAATCACTCAAAAGACCTTTGAA 1980
Db      1921 GTTTAACTGACGATATGACAGAGATCCCGAAATCACTCAAAAGACCTTTGAA 1980
Qy      1981 ATCAACCAACCAATACACTTACATCAAAAGATCAATGATCTTATTTGGGCTT 2040
Db      1981 ATCAACCAACCAATACACTTACATCAAAAGATCAATGATCTTATTTGGGCTT 2040
Qy      2041 AAACAGGAGTAAACAATCTTAAACGCTTCTTAATGGGAGTGTCTTCCAGCCAG 2100
Db      2041 AAACAGGAGTAAACAATCTTAAACGCTTCTTAATGGGAGTGTCTTCCAGCCAG 2100
Qy      2101 CCTACCTGTGTGTACACAGCAGAGACATGATTCAGTCAACAGCATACAGCTGT 2160
Db      2101 CCTACCTGTGTGTACACAGCAGAGACATGATTCAGTCAACAGCATACAGCTGT 2160
Qy      2161 CTGAAGAAACGTTCTTCAACAGCTTGAATCAAAATGTTTACTTAAATGAATTA 2220
Db      2161 CTGAAGAAACGTTCTTCAACAGCTTGAATCAAAATGTTTACTTAAATGAATTA 2220
```


OY		2221	GAGACTACCTCAGCCCTTAAATGCTTTTAATCAATAAATCTGGAAAGCTAGACTGAACCA	2280
Dd		2221	AGACTACTTCAGCCCTTAAATGCTTTTAATCAATAAATCTGGAAAGCTAGACTGAACCA	2280
OY		2281	TTGGAAACATTTAACTCAGACTCTGGATTCAAGAGTGCGGAAACCCCTTAGTCTCATGTGAAT	2340
Dd		2281	TTGGAAACATTTAACTCAGACTCTGGATTCAAGAGTGCGGAAACCCCTTAGTCTCATGTGAAT	2340
OY		2341	CCAAAGACAGCCACACCTTAGTATAGTATCTGCCCAAACCTAATGAGTTTAATAATCAAAATCT	2400
Dd		2341	CCAAGACAGCCACACCTTAGTATAGTATCTGCCCAAACCTAATGAGTTTAATAATCAAAATCT	2400
OY		2401	CCTTAAAAAATTT	2429
Dd		2401	CCTTAAAAAATTT	2429
RESULT 2				
AF057170				
LOCUS				
DEFINITION		AF057170	2420 bp mRNA linear PRI 17-OCT-1998	
ACCESSION		Homo sapiens dystrophin (VMD2) mRNA, alternatively spliced product,		
VERSION		complete cds.		
KEYWORDS		AF057170.1 GI:3335160		
SOURCE		.		
ORGANISM		Homo sapiens (human)		
REFERENCE		Homo sapiens		
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo. Petrushkin, K., Kolst, M. J., Bakell, B., Li, W., Xie, G., Marknell, T., Sandgren, O., Foreman, K., Holmgren, G., Andreasson, S., Vujic, M., Bergem, A. A., McGerty-Dugan, V., Figueroa, D., Austin, C. P., Metzker, M. L., Caskey, C. T. and Wadelius, C. Identification of the gene responsible for Best macular dystrophy Nat. Genet. 19 (3), 241-247 (1998) 98324772 MEDLINE PubMed 9662395 2 (bases 1 to 2420) Petrushkin, K. Direct Submision Submitted (03-APR-1998) Human Genetics, Merck Research Laboratories, West Point, PA 19486, USA		
FEATURES		Location/Qualifiers		
source		1..2420		
		/organism="Homo sapiens"		
		/mol_type="mRNA"		
		/db_xref="taxon:9606"		
		/chromosome="11"		
		/map="11q13"		
		1..2420		
		/gene="VMD2"		
		105..1412		
		/gene="VMD2"		
		/note="Best's macular dystrophy gene; BMD; alternatively splice isoform, contains exon 7"		
		/splicing_start=1		
		/product="dystrophin"		
		/protein_id="AAC64344.1"		
		/db_xref="GI:3335161"		
		/translation="MTITVTQSVANARLGSFRRLLICWRGRIYKLLIGEFILFLCYVY IIRFYRLALFEBOOLMEFKTLVCDSYLIOLIPISFVLGVTVLVTRMMWQYNLPML PDRLMSIVSGFGKDGKDGRGLLRRTILRYNALGNULILRSVTAHYKRPASOHLNOA GFNPFAHEKKDLSEKLSPHNMTWVPWPANLSKAMLAGR.RDPILOSILNENNTLR TGCHLAIAWDISTPLVTVQVVTAAVVSFFLTCLVGKOFNLPAAYGSHEDLVVFPVF TFPLPFVFGVKGLSGRALGMRHGQGHQQQLLETRMOQERKVASRESDQAMWPT PVIPATREAAGESSIEPGRRLIMWOSSSSTLEPRMMILIRPTGSTGICPCWLMMNR CTRRCLGSRCTGTGISPSHPSPDTQLPPSSVEPPLMAPPSISA"		
ORIGIN				
Query Match		99.6%	Score 2418.4;	DB 9; Length 2420;
Best Local Similarity		100.0%;	Pred. No. 0;	

	Matches	2419;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
Qy	1	CAGGGAGTCCCA	CGACCTAGTCG	CAGACCTTCTGTGGATCATCGGACCACCTGGAA	60					
Db	1	CAGGGAGTCCCA	CGACCTAGTCG	CAGACCTTCTGTGGATCATCGGACCACCTGGAA	60					
Qy	61	CCCCACCTGACCCMA	GGCCACCTGCTGAGGCCACCTGCTGGCCATGACATCACTTCA	120						
Db	61	CCCCACCTGACCCMA	GGCCACCTGCTGAGGCCACCTGCTGGCCATGACATCACTTCA	120						
Qy	121	CAAGCAAGTGGCTAAT	GCCTGCTTAAAGGCTCCTTCTCCGCTGCTGCTGTGCTGGCGGG	180						
Db	121	CAAGCAAGTGGCTAAT	GCCTGCTTAAAGGCTCCTTCTCCGCTGCTGCTGTGCTGGCGGG	180						
Qy	181	GCAGCATCTACAGAC	TGCTATATGCGAGTCTTAACTTTCCTGCTCTGCTACTACATCA	240						
Db	181	GCAGCATCTACAGAC	TGCTATATGCGAGTCTTAACTTTCCTGCTCTGCTACTACATCA	240						
Qy	241	TCGCGTTTATTATTA	TAGGCTGGCCCTTCACGGAAAGAACACACTGATGTTTGAAACCTGA	300						
Db	241	TCGCGTTTATTATTA	TAGGCTGGCCCTTCACGGAAAGAACACACTGATGTTTGAAACCTGA	300						
Qy	301	CTCTGTATTGGGACA	GCAGCTACATCCAGCTCATCCCCATTCCCTGCTGGGCTTCTACG	360						
Db	301	CTCTGTATTGGGACA	GCAGCTACATCCAGCTCATCCCCATTCCCTGCTGGGCTTCTACG	360						
Qy	361	TGACGCTGTCGTGAC	CCCGCTGTGTGGAAACAAGTACGAGAACCTGCGCTGGCCGACCGCC	420						
Db	361	TGACGCTGTCGTGAC	CCCGCTGTGTGGAAACAAGTACGAGAACCTGCGCTGGCCGACCGCC	420						
Qy	421	TCATGAGCTTGTGTG	TCGGGCTTTCGTCAAGGCAAGACGACGACGACGCGGCTGTGCGGC	480						
Db	421	TCATGAGCTTGTGTG	TCGGGCTTTCGTGTGCAAGGCAAGACGACGACGCGGCTGTGCGGC	480						
Qy	481	GCAGGCTCATCCGTA	GGCCAAACCTGGGCAACGTCATCTCTGGGCGGCTACGACACG	540						
Db	481	GCAGGCTCATCCGTA	GGCCAAACCTGGGCAACGTCATCTCTGGGCGGCTACGACACG	540						
Qy	541	CAGTCTACACACGCTT	CCCCAGCGCCACGACCTGTGTGACAGCAGGCTTTATGACTCCGG	600						
Db	541	CAGTCTACACACGCTT	CCCCAGCGCCACGACCTGTGTGACAGCAGGCTTTATGACTCCGG	600						
Qy	601	CAGAAACACAGACGAT	TGGAGAACTGAGCTTACACACACATGTTCTGGGTGCCCTGGG	660						
Db	601	CAGAAACACAGACGAT	TGGAGAACTGAGCTTACACACACATGTTCTGGGTGCCCTGGG	660						
Qy	661	TGTGTTTCCCACTGT	CAATGAAAGGCTGTGAGGCTGCAATCCGGGACCTTATCC	720						
Db	661	TGTGTTTCCCACTGT	CAATGAAAGGCTGTGAGGCTGCAATCCGGGACCTTATCC	720						
Qy	721	TGCTTCAGAGCTGCT	GTGAAAGAGTGAACCTTGTGCTGCACTGAGTGTGAACACTGTATG	780						
Db	721	TGCTTCAGAGCTGCT	GTGAAAGAGTGAACACTTGTGCTGCACTGAGTGTGAACACTGTATG	780						
Qy	781	CCTACGATGATTA	AGTATCCCACTGTGTATACACAGTGTGATCTGTGTGCGGTGTACA	840						
Db	781	CCTACGATGATTA	AGTATCCCACTGTGTATACACAGTGTGATCTGTGTGCGGTGTACA	840						
Qy	841	GCTTCTTCTGACTGT	TAGTGTGGCGGCAAGTTTCTGAAACCCAGCCCAAGGCTTACCTG	900						
Db	841	GCTTCTTCTGACTGT	TAGTGTGGCGGCAAGTTTCTGAAACCCAGCCCAAGGCTTACCTG	900						
Qy	901	GCATGAGCTGGA	CTCGTGTGCGCGCTTCACGTTCTGTCAAGTTCTTCTTATGTG	960						
Db	901	GCATGAGCTGGA	CTCGTGTGCGCGCTTCACGTTCTGTCAAGTTCTTCTTATGTG	960						
Qy	961	GCTGGTAGAAGTGG	CCCTTCCAGGGCCCTGTGGGCTGTGAGGCATGAGCCAGAGGGCTC	1020						
Db	961	GCTGGTAGAAGTGG	CCCTTCCAGGGCCCTGTGGGCTGTGAGGCATGAGCCAGAGGGCTC	1020						
Qy	1021	ATGGCAGACACTG	CTTGAGACGAGATGCACTGTGACAGAAAGAAAGTCTCAAGGGTATG	1080						
Db	1021	ATGGCAGACACTG	CTTGAGACGAGATGCACTGTGACAGAAAGAAAGTCTCAAGGGTATG	1080						

QY 1081 AAAGCAGCAGCGCTGTGTGCGCACAACCTGTAATCCAGACTACGCGGAGGCTGAGCAG 1140
 DB 1081 AAAGCAGCAGCGCTGTGTGCGCACAACCTGTAATCCAGACTACGCGGAGGCTGAGCAG 1140
 QY 1141 GAGAAATGGCTTGAACCCGCGAGGCGAGGTTGTGTGCGCAGAGAGCTCAATCAACCCCTT 1200
 DB 1141 GAGAAATGGCTTGAACCCGCGAGGCGAGGTTGTGTGCGCAGAGAGCTCAATCAACCCCTT 1200
 QY 1201 TGGAGAGGATGATGATATTTTGAACCACTGGATTTGTGAGAGAAATTTGAGAGTTC 1260
 DB 1201 TGGAGAGGATGATGATATTTTGAACCACTGGATTTGTGAGAGAAATTTGAGAGTTC 1260
 QY 1261 CTTGTTGGCTGTGTGATGATGACACGAGACTGCTCGGATGAGAGCCGGAATGTACTG 1320
 DB 1261 CTTGTTGGCTGTGTGATGATGACACGAGACTGCTCGGATGAGAGCCGGAATGTACTG 1320
 QY 1321 GAATTAAGCCGAGCCGACAGCCGCCCTTACACAGCTGCTCCGCCAGTTCCTGAGCCCTC 1380
 DB 1321 GAATTAAGCCGAGCCGACAGCCGCCCTTACACAGCTGCTCCGCCAGTTCCTGAGCCCTC 1380
 QY 1381 CTTTATGGGCTCCAGCTTCAACATCATGACCTGAACCAAGAGATGAGTTCCAGCCCA 1440
 DB 1381 CTTTATGGGCTCCAGCTTCAACATCATGACCTGAACCAAGAGATGAGTTCCAGCCCA 1440
 QY 1441 TCAGAGGAGCAGAGAGATGCTCAAGCTGAGCATGATGGCTGCTTCAAGGCTGAGTC 1500
 DB 1441 TCAGAGGAGCAGAGAGATGCTCAAGCTGAGCATGATGGCTGCTTCAAGGCTGAGTC 1500
 QY 1501 CCATGATCAACATCTCTCCAGGCGCAACTCAAGAGACCAATCTGTGCGCCCAAGAGGA 1560
 DB 1501 CCATGATCAACATCTCTCCAGGCGCAACTCAAGAGACCAATCTGTGCGCCCAAGAGGA 1560
 QY 1561 ATCCCTCTTCCACAGAGGCTCTGCCCAAAAACCAAGGAGCCCAACAGAACTTTAGGGG 1620
 DB 1561 ATCCCTCTTCCACAGAGGCTCTGCCCAAAAACCAAGGAGCCCAACAGAACTTTAGGGG 1620
 QY 1621 CCAGAGAACCAACAGAGGCTTGAAGCTTGAAGCTGAGCGCTTCAAGCTGAGCCCACT 1680
 DB 1621 CCAGAGAACCAACAGAGGCTTGAAGCTTGAAGCTGAGCGCTTCAAGCTGAGCCCACT 1680
 QY 1681 GTATCAGAGGCGAGGCTACTACAGTGCACCACAGAGCGCCCTCAGCGCCCACTCCATGTT 1740
 DB 1681 GTATCAGAGGCGAGGCTACTACAGTGCACCACAGAGCGCCCTCAGCGCCCACTCCATGTT 1740
 QY 1741 CTTCCCTTGAAGACCATCAGGCGCTCAAGCTTCAAGTGTCAAGGCAATGACACCA 1800
 DB 1741 CTTCCCTTGAAGACCATCAGGCGCTCAAGCTTCAAGTGTCAAGGCAATGACACCA 1800
 QY 1801 AGACAAAGCTTAAAGACTGTGAGTTCTGGGSCCAAGAAAGTTTGAATTTGCTCTAGA 1860
 DB 1801 AGACAAAGCTTAAAGACTGTGAGTTCTGGGSCCAAGAAAGTTTGAATTTGCTCTAGA 1860
 QY 1861 GAGCGATGGGCTTGTGATGAGACCCAGAAATCTTCAAGTGTGAGAGAGAAACTGTGGA 1920
 DB 1861 GAGCGATGGGCTTGTGATGAGACCCAGAAATCTTCAAGTGTGAGAGAGAAACTGTGGA 1920
 QY 1921 GTTTAACTGACGATATGACAGAGATCCCGAAATCACTCTCAAGAAACCTTTGGAACA 1980
 DB 1921 GTTTAACTGACGATATGACAGAGATCCCGAAATCACTCTCAAGAAACCTTTGGAACA 1980
 QY 1981 ATCAACCAACCAATACACTCACTCAAGATCATGAGATCTTAATTTGGGCTTGA 2040
 DB 1981 ATCAACCAACCAATACACTCACTCAAGATCATGAGATCTTAATTTGGGCTTGA 2040
 QY 2041 AAAAGAGATGAGACCAATCTTAACTGCTTCTTAATTTGGGCTTGA 2100
 DB 2041 AAAAGAGATGAGACCAATCTTAACTGCTTCTTAATTTGGGCTTGA 2100
 QY 2101 CCTCACTGTGTGTACACAGAGAGACTGATCCAGTCAACAGCCATACAGCTGTTCACA 2160
 DB 2101 CCTCACTGTGTGTACACAGAGAGACTGATCCAGTCAACAGCCATACAGCTGTTCACA 2160

QY 2161 CTGAAGAACGTGTCTCTACACAGCCTGATCAAAATGTTACTTAAATGATAAAAATCCC 2220
 DB 2161 CTGAAGAACGTGTCTCTACACAGCCTGATCAAAATGTTACTTAAATGATAAAAATCCC 2220
 QY 2221 AGACTACTTCAAGCCTTTTAAATGCTTTTAAATCATAAAAATCTGAAACCTAGACTGAACCA 2280
 DB 2221 AGACTACTTCAAGCCTTTTAAATGCTTTTAAATCATAAAAATCTGAAACCTAGACTGAACCA 2280
 QY 2281 TTGAAACATTTAACTCAGACTGTGATTCAGAGTCCGGAACCTTGTGTTCTATCTGAAT 2340
 DB 2281 TTGAAACATTTAACTCAGACTGTGATTCAGAGTCCGGAACCTTGTGTTCTATCTGAAT 2340
 QY 2341 CCAAGACAGCCACCTTATGATTTCTGCCCAACTAATGATTTAATTAATCAAAATACT 2400
 DB 2341 CCAAGACAGCCACCTTATGATTTCTGCCCAACTAATGATTTAATTAATCAAAATACT 2400
 QY 2401 CGTTAAAAAATTTTTAAAAA 2420
 DB 2401 CGTTAAAAAATTTTTAAAAA 2420
 RESULT 3
 BD136720 2229 bp DNA linear PART 18-SRP-2002
 LOCUS Best's macular dystrophy gene.
 DEFINITION BD136720 1 GI:23231665
 ACCESSION BD136720 1 GI:23231665
 VERSION BD136720 1 GI:23231665
 KEYWORDS JP 2002504559-A/2.
 SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS Petrukhin, K., Caskey, T. C., Metzker, M. and Wadelius, C.
 TITLE Best's macular dystrophy gene
 JOURNAL Patent: JP 2002504559-A 2 12-FEB-2002;
 COMMENT
 OS Homo sapiens (human)
 PN JP 2002504559-A/2
 PD 12-FEB-2002
 PR 22-FEB-1999 JP 2000533447
 PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI
 KONSTANTIN PETRUKHIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI
 WADELIUS
 PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC
 C12P21/09,
 CC C12N5/00, C12N15/00
 FH Best's macular dystrophy gene
 FT source Location/Qualifiers
 FT 1.2229
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 FEATURES
 source Location/Qualifiers
 1.2229
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 82.9%; Score 2013; DB 6; Length 2229;
 Best Local Similarity 91.6%; Pred. No. 0;
 Matches 2226; Conservative 0; Mismatches 0; Indels 203; Gaps 1;
 QY 1 CAGGAGTCCACAGAGCTTGTGCGCAGACCTTGTGTGAGATCATCGGACCCACCTGAA 60
 DB 1 CAGGAGTCCACAGAGCTTGTGCGCAGACCTTGTGTGAGATCATCGGACCCACCTGAA 60
 QY 61 CCCACCTGAGCCCAAGCCCACTGCTGAGGCCACCTGCTGGGCCATGACATCATCTTACA 120
 DB 61 CCCACCTGAGCCCAAGCCCACTGCTGAGGCCACCTGCTGGGCCATGACATCATCTTACA 120
 QY 121 CAAAGCAAGTGTGATAGCCGCTTAGGCTCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTG 180
 DB 121 CAAAGCAAGTGTGATAGCCGCTTAGGCTCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTG 180

QY	181	GCAGATCTACAGAGCTGCTATATGCGAGAGTTCTTAATCTTCTGCTGCTGCTACTACATCA	240
Db	181	GCAGATCTACAGAGCTGCTATATGCGAGAGTTCTTAATCTTCTGCTGCTGCTACTACATCA	240
QY	241	TCGCGTTATTTATAGAGCTGCGCCTTCA CGSAGAAACAACAGCTGATGTTTGAAACTGA	300
Db	241	TCGCGTTATTTATAGAGCTGCGCCTTCA CGSAGAAACAACAGCTGATGTTTGAAACTGA	300
QY	301	CTCTGATTTGCGACAGCTACATCCAGCTCATCCCATTTCTTCTGCTGGCTTCTACG	360
Db	301	CTCTGATTTGCGACAGCTACATCCAGCTCATCCCATTTCTTCTGCTGGCTTCTACG	360
QY	361	TGAGCGCTGGTGTGACCCGCTGGTGGAAACAATTAAGAGAACTGGCCGTGGCCCGACCGC	420
Db	361	TGAGCGCTGGTGTGACCCGCTGGTGGAAACAATTAAGAGAACTGGCCGTGGCCCGACCGC	420
QY	421	TCATAGAGCTGGTGTGCGGCTTTCGTGAAGCGAAGAACGAGCAAGGCTCGCTGCGGC	480
Db	421	TCATAGAGCTGGTGTGCGGCTTTCGTGAAGCGAAGAACGAGCAAGGCTCGCTGCGGC	480
QY	481	GCACGCTCATCCGCTACGCGCAACTGGCGAAGTGCTCATCTCGCGCAAGCTCAACACCG	540
Db	481	GCACGCTCATCCGCTACGCGCAACTGGCGAAGTGCTCATCTCGCGCAAGCTCAACACCG	540
QY	541	CAGTTACAAAGGCTTCCCGACGCGCCAGCACTGCTGCAACACGCTTATAGACTCCGG	600
Db	541	CAGTTACAAAGGCTTCCCGACGCGCCAGCACTGCTGCAACACGCTTATAGACTCCGG	600
QY	601	CAGAACACAAAGCACTTGGAGAACTGAGCTTACACACAACTGTTCTTGGTCCCTGGG	660
Db	601	CAGAACACAAAGCACTTGGAGAACTGAGCTTACACACAACTGTTCTTGGTCCCTGGG	660
QY	661	TGTGTTTTCGCAACTGTTCATGAAGGCGTGGCTTGGAGTTCGAATCCGAGCCCTATCC	720
Db	661	TGTGTTTTCGCAACTGTTCATGAAGGCGTGGCTTGGAGTTCGAATCCGAGCCCTATCC	720
QY	721	TGCTCCAGAGCGCTGTGAAACGAGATGAACAACCTTGCGTACTCAGTGTGACACCTGTATG	780
Db	721	TGCTCCAGAGCGCTGTGAAACGAGATGAACAACCTTGCGTACTCAGTGTGACACCTGTATG	780
QY	781	CCTACGACTGTGATTTAGTATCCCACTGTGTATATACAAGGTGTGACTGTGGCGGTATCA	840
Db	781	CCTACGACTGTGATTTAGTATCCCACTGTGTATATACAAGGTGTGACTGTGGCGGTATCA	840
QY	841	GCTTCTTCTGACTGTGTCTAGTTGGGCGGCACTTCTGTAACCCACAGGCTTACCCCTG	900
Db	841	GCTTCTTCTGACTGTGTCTAGTTGGGCGGCACTTCTGTAACCCACAGGCTTACCCCTG	900
QY	901	GCATAGAGCTGGACCTCGTGTGACCCTTCAAGTCCCTGAGTTCCTTCTTAATGTTG	960
Db	901	GCATAGAGCTGGACCTCGTGTGACCCTTCAAGTCCCTGAGTTCCTTCTTAATGTTG	960
QY	961	GCTGGCTGAAGGTGGGCTTCTTCAGAGGCGCTGCTGGGCTGAGGCAATGCCAGAGGGGTC	1020
Db	961	GCTGGCTGAAGGTGGGCTTCTTCAGAGGCGCTGCTGGGCTGAGGCAATGCCAGAGGGGTC	1020
QY	1021	ATGGCGACAGACTGCTTGAAGACGAGAGATGCAAGTGTCAAGAAAGAAAGTCTCA CGGTTAG	1080
Db	971	-----	970
QY	1081	AAAGCAGCAGGCGTGTGTGGCGCACACTGTAACTCCAGCTACTCGGACAGCTGAGGAGC	1140
Db	971	-----	970
QY	1141	GAGAAATCGCTTGAAACCCGGAGGCGGAGGTTGTGTGGCAGAGCACTTCAATCAACCCCTT	1200
Db	971	-----GGTGGCAGAGCAAGCTCATCAACCCCTT	997
QY	1201	TGAGAGAGATGATGATTTTGTGAACCAACTGATTTGTTCGACAGAGATTTGCAGAGTGTG	1260
Db	998	TGAGAGAGATGATGATTTTGTGAACCAACTGATTTGTTCGACAGAGATTTGCAGAGTGTG	105

QY	1261	CCTGTGGCTCTGTGATGATGATGACACAGGACCTGCTCGGATGTGAGCCCGGACATGTACTG	1322
Db	1058	CCGTGTGGCTGTGTGATGATGATGACACAGGACCTGCTCGGATGTGAGCCCGGACATGTACTG	1117
QY	1321	GAATTAAGCCCGAGCCACAGCCCCCTTACACAGCTGCTTCCGCCAGATTCCGTTCGAGCTTC	1380
Db	1118	GAATTAAGCCCGAGCCACAGCCCCCTTACACAGCTGCTTCCGCCAGATTCCGTTCGAGCTTC	1177
QY	1391	CTTTATGGGCTCCACCTTTCAACTCAGCTGAAACAAAGAGATGAGATTCCAGCCCA	1440
Db	1178	CTTTATGGGCTCCACCTTTCAACTCAGCTGAAACAAAGAGATGAGATTCCAGCCCA	1237
QY	1441	TCAGAGAGACGAGAGAGATGCTCAGCGCTGGATCATTTGGCCGCTTCTTAAGGCTTCGACTC	1500
Db	1238	TCAGAGAGACGAGAGAGATGCTCAGCGCTGGATCATTTGGCCGCTTCTTAAGGCTTCGACTC	1297
QY	1501	CCATGATCAGCATCTCTCCAGGGCAAACTCAAGGACCAAACTACTGTGTGGCCCAAGAGGA	1560
Db	1298	CCATGATCAGCATCTCTCCAGGGCAAACTCAAGGACCAAACTACTGTGTGGCCCAAGAGGA	1355
QY	1561	ATCCCTTCTCCAGGAGGCTGTCCCAAAAACAAGGACGACCAAGACGATTAGGGG	1620
Db	1358	ATCCCTTCTCCAGGAGGCTGTCCCAAAAACAAGGACGACCAAGACGATTAGGGG	1411
QY	1621	CCAGGAGACAAACAAGGCTTGGAAGCTTAAAGCTGTGGAGCGCTTCAAGTCTGGCCCACT	1688
Db	1418	CCAGGAGACAAACAAGGCTTGGAAGCTTAAAGCTGTGGAGCGCTTCAAGTCTGGCCCACT	1477
QY	1681	GTATCAGAGGCGAGGCTCTACAGTGGCCCAACAGCGCCCTCAGGCCCACTCCCACTTT	1744
Db	1478	GTATCAGAGGCGAGGCTCTACAGTGGCCCAACAGCGCCCTCAGGCCCACTCCCACTTT	1537
QY	1741	CTTCCCCCTTAACAACCATCAGCGCGCTCAAAAGCTTCAACAGTGTCAAGGATAGACACAA	1800
Db	1538	CTTCCCCCTTAACAACCATCAGCGCGCTCAAAAGCTTCAACAGTGTCAAGGATAGACACAA	1597
QY	1801	AGACAAAAGCTTAAAGACTGTGAGTCTGTGGGCGCAAGAAAAGTTTGAATTGCTCTCAGA	1860
Db	1598	AGACAAAAGCTTAAAGACTGTGAGTCTGTGGGCGCAAGAAAAGTTTGAATTGCTCTCAGA	1657
QY	1861	GAGGATGGGGCCCTTGATGAGGACCCAGAAAGTCTTCAAGTGGAGGAGAAAACGTGTGA	1922
Db	1658	GAGGATGGGGCCCTTGATGAGGACCCAGAAAGTCTTCAAGTGGAGGAGAAAACGTGTGA	1717
QY	1921	GTTTAACTCTGAGATATGCGAGAGATCCCGGAAAATCACTCAAGAATCCTTTGGAACA	1980
Db	1718	GTTTAACTCTGAGATATGCGAGAGATCCCGGAAAATCACTCAAGAATCCTTTGGAACA	1777
QY	1981	ATCACCACCAACAATACACTACACTCAAGATCAATGATCTCTTATTTGGGCTTTGGA	2040
Db	1778	ATCACCACCAACAATACACTACACTCAAGATCAATGATCTCTTATTTGGGCTTTGGA	1837
QY	2041	AAACAGGAGATGAAGCACTTCTTAACTGTCTTCTTAATGGGATGCTTCCGCAAGCAGGT	2100
Db	1838	AAACAGGAGATGAAGCACTTCTTAACTGTCTTCTTAATGGGATGCTTCCGCAAGCAGGT	1897
QY	2101	CCTCAGCCTGTGTACACACAGCAGAGACCTGATCAGTCAAGCAGCATACAGCTGTCCACA	2166
Db	1898	CCTCAGCCTGTGTGTACACACAGCAGAGACCTGATCAGTCAAGCAGCATACAGCTGTCCACA	1955
QY	2161	CTGAAGAACGTGTCTCTCAACAGGCTGATCAAAATGTTAGCTTAAATGATPAAAAATCC	2220
Db	1958	CTGAAGAACGTGTCTCTCAACAGGCTGATCAAAATGTTAGCTTAAATGATPAAAAATCC	2017
QY	2221	AGACTACTTTCAGGCTTTTAAATGCTTTTAAATGATPAAAAATGTAAGCTGAACCA	2280
Db	2018	AGACTACTTTCAGGCTTTTAAATGCTTTTAAATGATPAAAAATGTAAGCTGAACCA	2077
QY	2281	TTGGAAACATTTAACTCAGACTCTGGATTCAGATCGGAAACCTTATGTTCTAATCTGAAT	2340
Db	2078	TTGGAAACATTTAACTCAGACTCTGGATTCAGATCGGAAACCTTATGTTCTAATCTGAAT	2137
QY	2341	CCAAGACAGCCACCTTAGTATCTGCGCCAACTAATGAGTTTAAATPACAAATACT	2400

Db 2138 CCAGAGAGCCACCTTAGTATCTGCCCAACTATGAGTTTAAATTAATAACT 2197
Oy 2401 CGTTAAAAAAGAAAAAAGAAAAA 2429
Db 2198 CGTTAAAAAAGAAAAAAGAAAAA 2226

RESULT 4
AF057169 2210 bp mRNA linear pri 17-OCT-1998
LOCUS Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product,
complete cds.
AF057169
VERSION AF057169.1 GI:3335158
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2210)
Petrushkin, K., Koishi, M.J., Bakali, B., Li, W., Xie, G., Marknell, T.,
Sander, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, W.,
Metzger, M.L., Caskey, C.T., and Wadelius, C.
Identification of the gene responsible for Best macular dystrophy
Nat. Genet. 19 (3), 241-247 (1998)
98324772
MEDLINE 2 (bases 1 to 2210)
Petrushkin, K.
REFERENCE Direct Submission
AUTHORS Submitted (03-APR-1998) Human Genetics, Merck Research
JOURNAL Laboratories, West Point, PA 19486, USA
FEATURES
source 1..2210
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
1..2210
/gene="VMD2"
105..1862
/gene="VMD2"
/note="Best's macular dystrophy gene; BMD; alternatively
splice isoform"
/codon_start=1
/product="bestrophin"
/protein_id="AAC6434.1"
/db_xref="GI:3335159"
/translation="MTITTYTSQVANARLGSFSLLLCMRGSIVKLYGFLLFLCY
IIRIYRLALTEQOLMEKLTLYCDSYIOLIPISFVLAGVYTLVTRMNOYENIPW
PDRMLSVSGEKGDEQGLRLRLTLRYANLGNVLIRSVTAAYKRPASQHLVQA
GFMTPAHKGLEKSLPHEMFWVWFAVLSMKAWLGGIRIDPILLOSILNEMTLR
TOGHLAYAWMISILVYTVTVAVYSFLTCLVGRQFNPKALFGLHLVAVF
TFLOPFYVGLKLYAEOLINPFGEDDDEFETNMIVDKILQVSLAVDEHQDLPRNEP
DMYWKRPDPPTTAAQAQFRASFWKSTENISLNKEEMFOPNOBEBDAAGIIGR
FLGIQSHDHPPRANSRTKLIWPKBSLHEGLPKNHKAKQVNRQOENDKAKLAV
DAFKAAPLYORPGYSAQPTPLSPTPMFPLEPSASPKLSVIGIDTKKSLKTVSSG
AKSFELLESDEGLMEHPVSQVRKTVFNLIDMPEIPENHLKRPIDQSPTNIHTT
LKDHMDPYMALERDEAHS"

ORIGIN
Query Match 82.1%; Score 1995.4; DB 9; Length 2210;
Best Local Similarity 91.5%; Pred. No. 0;
Matches 2209; Conservative 0; Mismatches 1; Indels 203; Gaps 1;

Oy 1 CAGGAGATCCACAGCCTAGCTGCGACAGCTTCTGTGGATCATGGACCCAGCTGGA 60
Db 1 CAGGAGATCCACAGCCTAGCTGCGACAGCTTCTGTGGATCATGGACCCAGCTGGA 60
Oy 61 CCCACCTGACCCAGAGCCACCTGCTGACGCCACCTGCGCATGACCATCACTTACA 120

Db 61 CCCACCTGACCCAGAGCCACCTGCTGACGCCACCTGCGCATGACCATCACTTACA 120
Oy 121 CAGGCAAGTGGCTAATGCGCGCTTAGGCTCTTCGCGCTGCTGCTGCTGCGGG 180
Db 121 CAGGCAAGTGGCTAATGCGCGCTTAGGCTCTTCGCGCTGCTGCTGCTGCGGG 180
Oy 181 GAGCATCTTACAGCTGCTATATGAGCGAGTCTTAACTCTTCTGCTGCTACTACATCA 240
Db 181 GAGCATCTTACAGCTGCTATATGAGCGAGTCTTAACTCTTCTGCTGCTACTACATCA 240
Oy 241 TCGGCTTATTTATAGGCTGCGCGCTGCGCGAGAAACAGCTGATGTTGAAGAACTGA 300
Db 241 TCGGCTTATTTATAGGCTGCGCGCTGCGCGAGAAACAGCTGATGTTGAAGAACTGA 300
Oy 301 CTGTATATGAGCAGCTGATCATGCTCATCCCATTTCTTCTGCTGCTGCTGCTGCTG 360
Db 301 CTGTATATGAGCAGCTGATCATGCTCATCCCATTTCTTCTGCTGCTGCTGCTGCTG 360
Oy 361 TGAGCTGCTGCTGAGCCCGCTGCTGAGAACAGTACAGAACTGCGCTGCGCGAGCC 420
Db 361 TGAGCTGCTGCTGAGCCCGCTGCTGAGAACAGTACAGAACTGCGCTGCGCGAGCC 420
Oy 421 TCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Oy 481 GCAAGCTCATCGCTACGCCCAACCTGGGCAAGTGTCTATCTGCGCAGCTGACAGCG 540
Db 481 GCAAGCTCATCGCTACGCCCAACCTGGGCAAGTGTCTATCTGCGCAGCTGACAGCG 540
Oy 541 CAGTCTTACAGGCGCTTCCCAAGCGCCACACCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 CAGTCTTACAGGCGCTTCCCAAGCGCCACACCTGCTGCTGCTGCTGCTGCTGCTG 600
Oy 601 CAGAACACAGAGCTTGAAGAACTGAGCTTACCAACAATGTTCTGAGGCTTACCTGG 660
Db 601 CAGAACACAGAGCTTGAAGAACTGAGCTTACCAACAATGTTCTGAGGCTTACCTGG 660
Oy 661 TGTGTTTGGCAACCTGTATATGAAGCGGTGCTGAGAGTGAATCCGGGACCTATCC 720
Db 661 TGTGTTTGGCAACCTGTATATGAAGCGGTGCTGAGAGTGAATCCGGGACCTATCC 720
Oy 721 TGTCTCAGAGCGCTGCTGAGAGAGATGAACACTTGGCTGATCACTGATGATG 780
Db 721 TGTCTCAGAGCGCTGCTGAGAGAGATGAACACTTGGCTGATCACTGATGATG 780
Oy 781 CCTACGACTGATATGATATCCCACTGCTGATATCAAGAGTGTGCTGCGGTATACA 840
Db 781 CCTACGACTGATATGATATCCCACTGCTGATATCAAGAGTGTGCTGCGGTATACA 840
Oy 841 GCTTCTTCTCTGATCTGCTGATGAGGCGGAGTTTCTGAACCCAGGACCAAGCTTACCCTG 900
Db 841 GCTTCTTCTCTGATCTGCTGATGAGGCGGAGTTTCTGAACCCAGGACCAAGCTTACCCTG 900
Oy 901 GCCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 GCCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Oy 961 GCTGCTTAAAGTGGGCTTCTTCAAGGCGCTGCTGAGGCTGAGGCAAGAGGGGTC 1020
Db 961 GCTGCTTAAAGTGGGCTTCTTCAAGGCGCTGCTGAGGCTGAGGCAAGAGGGGTC 1020
Oy 1021 ATGGCAGACAGCTGTTGAGACGAGAGTGCAGTGTACAGAAAGAGAGTCTCACGGGTAG 1080
Db 971 ----- 970
Oy 1081 AAAGCAGCAGGCGTGTGCGGCAACCTGTAAATCCAGCTACTCGGAGAGCTGAGGAG 1140
Db 971 ----- 970
Oy 1141 GAGAAATGCTTGAACCCGGAGGCGGAGGTTGTGTGTGCGAGAGAGCTATCAACCCCTT 1200
Db 971 ----- 970

```

Db      971 -----GATGGCAGAGCAGCTCATCAACCCCTT 997
Qy      1201 TTGAGAGGATGATGATGATTTTGAGACCAACTGGATTGTGAGAGGAATTTGAGGTGTC 1260
Db      998 TGGAGAGGATGATGATGATTTTGAGACCAACTGGATTGTGAGAGGAATTTGAGGTGTC 1057
Qy      1261 CTTGTGTGCTGTGATGATGATGACACAGACCTGCTCGATGATGAGCCGACATGTAATG 1320
Db      1058 CTTGTGTGCTGTGATGATGATGACACAGACCTGCTCGATGATGAGCCGACATGTAATG 1117
Qy      1321 GAATTAAGCCGAGACCAAGCCCTTACACAGCTGCTTCGCGCAATTCGTCAGGCTC 1380
Db      1118 GAATTAAGCCGAGACCAAGCCCTTACACAGCTGCTTCGCGCAATTCGTCAGGCTC 1177
Qy      1381 CTTTATGGGCTCCACTTCAACATCAGCTGTAACAAAGAGGATGAGATTCGAGCCCA 1440
Db      1178 CTTTATGGGCTCCACTTCAACATCAGCTGTAACAAAGAGGATGAGATTCGAGCCCA 1237
Qy      1441 TCAAGAGAGCAGAGAGGATGCTCAGCTGATCATTTGGCCGCTTCTAGGCTGCAATC 1500
Db      1238 TCAAGAGAGCAGAGAGGATGCTCAGCTGATCATTTGGCCGCTTCTAGGCTGCAATC 1297
Qy      1501 CCATGATCACCATTCTCCAGAGGCAAACTCAAGACCAAACTACTGTGACCCAAAGGGA 1560
Db      1298 CCATGATCACCATTCTCCAGAGGCAAACTCAAGACCAAACTACTGTGACCCAAAGGGA 1357
Qy      1561 ATCCCTTCTCCAGAGGCTGCGCAAAACCAAGAGCCCAACCAAGAGGAGG 1620
Db      1358 ATCCCTTCTCCAGAGGCTGCGCAAAACCAAGAGCCCAACCAAGAGGAGG 1417
Qy      1621 CCAAGAGAGCAACAAGGCTGGAAGCTGAGGACGCTTCAAGTCTGAGCCCACT 1680
Db      1418 CCAAGAGAGCAACAAGGCTGGAAGCTGAGGACGCTTCAAGTCTGAGCCCACT 1477
Qy      1681 GATCAGAGGCGAGGCTACTACAGTCCCAAGAGGCTTCAAGTCTGAGGAGG 1740
Db      1478 GATCAGAGGCGAGGCTACTACAGTCCCAAGAGGCTTCAAGTCTGAGGAGG 1537
Qy      1741 CTTCCCTTGAACCATCAGGCGCTGAAAGCTTCAAGTCTCAGAGGATGACACCA 1800
Db      1538 CTTCCCTTGAACCATCAGGCGCTGAAAGCTTCAAGTCTCAGAGGATGACACCA 1597
Qy      1801 AGACAAAGCTTAAAGACTGATGATGAGGAGCCCAAGAAAGTTTGAATGCTCTGAG 1860
Db      1598 AGACAAAGCTTAAAGACTGATGATGAGGAGCCCAAGAAAGTTTGAATGCTCTGAG 1657
Qy      1861 GAGGATGAGGAGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAACTGTGGA 1920
Db      1658 GAGGATGAGGAGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAAGAGAAACTGTGGA 1717
Qy      1921 GTTTAACCTGACGATATGCGAGAGATCCCGAAATCACTCAAAAGACCTTTGGAACA 1980
Db      1718 GTTTAACCTGACGATATGCGAGAGATCCCGAAATCACTCAAAAGACCTTTGGAACA 1777
Qy      1981 ATCAGCAACCAATACACTACACTCAAAAGATCAGATGATGATGAGGAGCTTGA 2040
Db      1778 ATCAGCAACCAATACACTACACTCAAAAGATCAGATGATGATGAGGAGCTTGA 1837
Qy      2041 AAACAGGAGTGAAGACATTCCTTAACCTGCTTCTTAATGGGAGTCTTCGCGACCAAGT 2100
Db      1838 AAACAGGAGTGAAGACATTCCTTAACCTGCTTCTTAATGGGAGTCTTCGCGACCAAGT 1897
Qy      2101 CCTTACCTGTGTGATGACACAGAGACATGATGATGATGATGATGATGATGATGATGAT 2160
Db      1898 CCTTACCTGTGTGATGACACAGAGACATGATGATGATGATGATGATGATGATGATGAT 1957
Qy      2161 CTGAAGAAAGCTGTCTCAACAGGCTGAATCAAAATGTTAGCTTAATAGATTAATATCC 2220
Db      1958 CTGAAGAAAGCTGTCTCAACAGGCTGAATCAAAATGTTAGCTTAATAGATTAATATCC 2017
Qy      2221 AGACTACTTACGCTTTAATGCTTTTATGATTAATAAATCTGTGAAGCTAGATGAACCA 2280
Db      2018 AGACTACTTACGCTTTAATGCTTTTATGATTAATAAATCTGTGAAGCTAGATGAACCA 2077

```

```

Qy      2281 TTGAAACATTTAACTCAGACTGTGATGAGAGTCCGGAACCTTAGTCTATCTGAT 2340
Db      2078 TTGAAACATTTAACTCAGACTGTGATGAGAGTCCGGAACCTTAGTCTATCTGAT 2137
Qy      2341 CCAAGACGACCACTTAGTATCTGCCCAACTAATGATTTAATTAATCAAAATCT 2400
Db      2138 CCAAGACGACCACTTAGTATCTGCCCAACTAATGATTTAATTAATCAAAATCT 2197
Qy      2401 CGTTAAAAAAA 2413
Db      2198 CGTTAAAAAAA 2210

```

RESULT 5

AF073501

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PubMed

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

gene

CDS

ORIGIN

Query Match

Best Local Similarity

Matches 2168; Conservative

80.4%; Score 1953.2; DB 9; Length 2171;

Pred. No. 0;

Mismatches 3; Indels 203; Gaps 1;

AKSFELLSDESDDALMEHREVSQVRKTVFNLDMBEIPENHLKEPLEDSPTNIHTT

LKHMDPYWALKEHREDAHS"

/product="vitellogenin macular dystrophy protein"

/protein_id="AAC3766.1"

/db_xref="GI:3511242"

/translation="MTITTSQVANAARLGSFSLRLCWRGSYKLYGEPLIFLCY

IRFIRLALTEBQDLMEFLTLTCDYSLIPLISFLGFTVLVTRMNOENLPL

PRLMSLVSGFEGKDEQRLRLTLRLVNLGVLIRSVSTAVYRPSAOHLVDA

GPMTPRBNKOLEKLSLPHNFWPMVWFANLSMKAWLGRIRDEILQSLNENYLR

TDCGGLVAYDWISILPLVYQVTVAVSFLTCLVGRQPLPAKAYGHELDLVVP

TPIQPFVYWMKLYABGLNPRGDDDPFTNIVPNNLVSLAUPBMDLVPMPR

DMYMKRBPPTTASAPRRASFMKSTNLSLNKSEMEFQPNQDEBDNAQITGR

FLQIASHDHNRPRANSRTKLWPKRESLLHEGLPKHKAQKQNGQEDKAWLKLAV

DAFKSAFLYORPGVSAPOPTLSPTPWFPELPBSAPSKLSVTGIDTKDLSLKVSG

AKSFELLSDESDDALMEHREVSQVRKTVFNLDMBEIPENHLKEPLEDSPTNIHTT

LKHMDPYWALKEHREDAHS"

ORIGIN

OY	17	CCTAATGCGCAACCTTCTGTGGAACTCATCGAACCCACTTGGAACCCCACTGACCAG	76
Db	1	CTTAATGCGCAACCTTCTGTGGAACTCATCGAACCCACTTGGAACCCCACTGACCAG	60
OY	77	CCCACCTGTGCAGGCCACTGCGCTGGGCCATGACATCATTTCACAAGCGMAATGGCTTA	13
Db	61	CCCACCTGTGCAGGCCACTGCGCTGGGCCATGACATCATTTCACAAGCGMAATGGCTTA	12
OY	137	TGCCCGCTTAGGCTCTTTCTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	19
Db	121	TGCCCGCTTAGGCTCTTTCTCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	18
OY	197	GCTATATAGGAGGTTCTTAATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	25
Db	181	GCTATATAGGAGGTTCTTAATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	24
OY	257	GCTGCGCTTCAACGAGAACAACAGCTGATGTGTGAGAACTGACTGTATATGACAG	31
Db	241	GCTGCGCTTCAACGAGAACAACAGCTGATGTGTGAGAACTGACTGTATATGACAG	30
OY	317	CTACATCCAGCTCATCCCCATTTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	37
Db	301	CTACATCCAGCTCATCCCCATTTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	36
OY	377	CCGCTGTGAAACAGTACGAGAACTGTGCGTGTGCGAACCGGCTTCAATGAGCTGTGTG	43
Db	361	CCGCTGTGAAACAGTACGAGAACTGTGCGTGTGCGAACCGGCTTCAATGAGCTGTGTG	42
OY	437	GGGCTTGTGGAAGGCAAGGACGAGCAAGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGT	49
Db	421	GGGCTTGTGGAAGGCAAGGACGAGCAAGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGT	48
OY	497	CGCCAACTGTGGCAACGTGTCTCATCTGTGCGAGGCTGTGTGTGTGTGTGTGTGTGTGT	55
Db	481	CGCCAACTGTGGCAACGTGTCTCATCTGTGCGAGGCTGTGTGTGTGTGTGTGTGTGTGT	54
OY	557	CCCCAGGCCCAAGCACTGTGTGCAAGCAAGGCTTTATGACTTGTGTGTGTGTGTGTGTGT	61
Db	541	CCCCAGGCCCAAGCACTGTGTGCAAGCAAGGCTTTATGACTTGTGTGTGTGTGTGTGTGT	60
OY	617	GGAGAACTGAGCCTTACACACAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	67
Db	601	GGAGAACTGAGCCTTACACACAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	66
OY	677	GTCATAAGAAAGCGCTGT	73
Db	661	GTCATAAGAAAGCGCGT	72
OY	737	GAAACGATGAAACCTTGT	79
Db	721	GAAACGATGAAACCTTGT	78
OY	797	TATCCACTGTGTGTATACACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	85
Db	781	TATCCACTGTGTGTATACACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	84
OY	857	TCTAGTTGGAGGCGATTTCTGAAACCCAGCCAGGCTTACCTGTGCCATGAGTGAAGCT	91
Db	841	TCTAGTTGGAGGCGATTTCTGAAACCCAGCCAGGCTTACCTGTGCCATGAGTGAAGCT	90
OY	917	CGTTGTGCGCGCTTCAAGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	97
Db	901	CGTTGTGCGCGCTTCAAGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	96
OY	977	CCTGTCCAGGCGCTGT	103
Db	955	-----	95
OY	1037	TGAGACGAGATGACAGTGTACAGAAAGAGAGTCTCACGGGTAAAGACAGCCAGCGTG	109
Db	955	-----	95

QY	1097	GTGGGGCAACACCTGTATATCCACACTCTCGGGAGGCTGAGCGAGAAATCGTTGAAC	1156
Db	955	-----	954
QY	1157	CGGAGGCGAGAGTGTGTGGCAGAGAGCTCATCAACCCCTTTGAGAGATGATAT	1216
Db	955	-----GGTGGCAGAGCGCTCATCAACCCCTTTGAGAGATGATAT	997
QY	1217	GATTTGAGCAACATGGAATTTGTCCACAGAAATTTTGCAAGTGTCCCTGTTGGCTGTGAT	1276
Db	998	GATTTGAGCAACATGGAATTTGTCCACAGAAATTTTGCAAGTGTCCCTGTTGGCTGTGAT	1057
QY	1277	GAGATGCAACGAGACCTGCTCGAATGAGCCGGAATGTATCTGAATTAAGCCGAGCCA	1336
Db	1058	GAGATGCAACGAGACCTGCTCGAATGAGCCGGAATGTATCTGAATTAAGCCGAGCCA	1117
QY	1337	CAGCCCCCTTACACAGCTGCTTCCGCCAGTTCCGTGAGCCCTCTTATAGGCTCCACC	1396
Db	1118	CAGCCCCCTTACACAGCTGCTTCCGCCAGTTCCGTGAGCCCTCTTATAGGCTCCACC	1177
QY	1397	TTCAATCATGAGCTGTAAACAAAGAGATGATTTCCAGCCCAATCAGAGAGACAGAG	1456
Db	1178	TTCAATCATGAGCTGTAAACAAAGAGATGATTTCCAGCCCAATCAGAGAGACAGAG	1237
QY	1457	GATGCTACGCTGGGATCATTTGGCCGCTTCTTGAAGCTGAGTCCCATGATCAACATCT	1516
Db	1238	GATGCTACGCTGGGATCATTTGGCCGCTTCTTGAAGCTGAGTCCCATGATCAACATCT	1297
QY	1517	CCAGAGGCAAACTCAAGAGCCAAACTATGTGGCCCAAGAGGATTCCTTCTCCAGAG	1576
Db	1298	CCAGAGGCAAACTCAAGAGCCAAACTATGTGGCCCAAGAGGATTCCTTCTCCAGAG	1357
QY	1577	GGCTGCCCCAAAACCAAGAGCCAAACAGAACGTTAGGGGCCAGAGAAACAAAG	1636
Db	1358	GGCTGCCCCAAAACCAAGAGCCAAACAGAACGTTAGGGGCCAGAGAAACAAAG	1417
QY	1637	GGCTGGAAGCTTAAGCGTGTGAGAGCCTTCAAGTCTGGCCCATGTATCAGAGCCAGGC	1696
Db	1418	GGCTGGAAGCTTAAGCGTGTGAGAGCCTTCAAGTCTGGCCCATGTATCAGAGCCAGGC	1477
QY	1697	TACTACAGTGCACCAAGAGCCCTCAGGCCCACTCCATGTTCTTCCCTTAAGACA	1756
Db	1478	TACTACAGTGCACCAAGAGCCCTCAGGCCCACTCCATGTTCTTCCCTTAAGACA	1537
QY	1757	TCAGGCGCGTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAGACAAAGCTTAAAG	1816
Db	1538	TCAGGCGCGTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAGACAAAGCTTAAAG	1597
QY	1817	ACTGTGAGTTCTGGGGCCAGAAAGTTTGAATGTCTCAGAGAGGATGGGGCCTTG	1876
Db	1598	ACTGTGAGTTCTGGGGCCAGAAAGTTTGAATGTCTCAGAGAGGATGGGGCCTTG	1657
QY	1877	ATGAGACACCCAGAGATCTCAAGTGAAGAGAAACGTGAGTTTAACTGACGAT	1936
Db	1658	ATGAGACACCCAGAGATCTCAAGTGAAGAGAAACGTGAGTTTAACTGACGAT	1717
QY	1937	ATGCAAGATTCGCCCAAAAAATCACTCCAAAGAACTTTTGAACAATCAACAACAATA	1996
Db	1718	ATGCAAGATTCGCCCAAAAAATCACTCCAAAGAACTTTTGAACAATCAACAACAATA	1777
QY	1997	CACACTACACTCAAAAGATCATGATCTCTTATTTGGGCTTGGAAAAACAGGATGAAGCA	2056
Db	1778	CACACTACACTCAAAAGATCATGATCTCTTATTTGGGCTTGGAAAAACAGGATGAAGCA	1837
QY	2057	CATTCCAAACCTGCTTCTTAATGGGAGTGTTCGCAAGCCAGGTCCTCACTGTGTAC	2116
Db	1838	CATTCCAAACCTGCTTCTTAATGGGAGTGTTCGCAAGCCAGGTCCTCACTGTGTAC	1897
QY	2117	ACCAGAGAGACATGATCAATGCAAGGCATACAGCTGTCCACATGAAGAAAGTGTCT	2176
Db	1898	ACCAGAGAGACATGATCAATGCAAGGCATACAGCTGTCTCCACATGAAGAAAGTGTCT	1957
QY	2177	ACAACAGCTGAATCAATGGTTAGCTTAATAGATTAATAATCCAGACTACTTCAGCTT	2236

Db 1958 ACAACAGCCTGAATCAATGTTAGCTTAATAGATAAAATCCAGACTACTGAGCCTT 2017
Qy 2237 TAAAGCCTTTATTCATATAAAACCTGTGAACCTAGACATGAACATTGAACATTAACT 2296
Db 2018 TAAAGCCTTTATTCATATAAAACCTGTGAACCTAGACATGAACATTGAACATTAACT 2077
Qy 2297 CAGACTGTGATTCAGATCGGGAACCTTAGTTCTATCTGTAATCAAGACGCCACAC 2356
Db 2078 CAGACTGTGATTCAGATCGGGAACCTTAGTTCTATCTGTAATCAAGACGCCACAC 2137
Qy 2357 TTAGTATACCTGCCCAACTAATGAGTTTAATAA 2390
Db 2138 TTAGTATACCTGCCCAACTAATGAGTTTAATAA 2171

RESULT 6
CQ729633 2170 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 15567 from Patent WO02068579.
DEFINITION CQ729633
ACCESSION CQ729633
VERSION CQ729633.1 GI:42301252
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W., and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A, 15567 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1. .2170
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 80.0%; Score 1942.8; DB 6; Length 2170;
Best Local Similarity 91.3%; Pred. No. 0;
Matches 2168; Conservative 0; Mismatches 2; Indels 204; Gaps 2;
Qy 17 CTTAGTGGCCAGACCTTGTGGGATCATCGAGCCACCTTGAAACCCCACTGACCAAG 76
Db 1 CCTAGTCGCCAGACCTTGTGGGATCATCGAGCCACCTTGAAACCCCACTGACCAAG 60
Qy 77 CCCACCTGCTGACGCCACCTGCGCCATGACCATCACTTACACAGCCAAAGTGGCTAA 136
Db 61 CCCACCTGCTGACGCCACCTGCGCCATGACCATCACTTACACAGCCAAAG-GGCTAA 119
Qy 137 TGCCCGCTTAGAGCTCTTCTCCGCGCTGCTGCTGCTGCGGGGACATCTCAAGCT 196
Db 120 TGCCCGCTTAGAGCTCTTCTCCGCGCTGCTGCTGCTGCGGGGACATCTCAAGCT 179
Qy 197 GCTAATAGCGAGTTCTTAATCTTCTGCTCTGCTACTACATCATCCGCTTATTATTAAG 256
Db 180 GCTAATAGCGAGTTCTTAATCTTCTGCTCTGCTACTACATCATCCGCTTATTATTAAG 239
Qy 257 GCTGAGCCTCAAGGAAGAAACAAGCTGATGTTTGAAGAACTGACTCTGTAATGGAAG 316
Db 240 GCTGAGCCTCAAGGAAGAAACAAGCTGATGTTTGAAGAACTGACTCTGTAATGGAAG 299
Qy 317 CTACATCAAGCTCATCCCATTTCTTGTGCTGCTGAGCTTCTACGTAAGCTGTGTAAC 376
Db 300 CTACATCAAGCTCATCCCATTTCTTGTGCTGCTGAGCTTCTACGTAAGCTGTGTAAC 359
Qy 377 CCGCTGTGTAACAAGTAAGAAACCTGCGGTGGCCGACCGCTCATGAGCTGTGTCTC 436
Db 360 CCGCTGTGTAACAAGTAAGAAACCTGCGGTGGCCGACCGCTCATGAGCTGTGTCTC 419

Qy 437 GGGCTTCGTGAAAGCAAGACGAGCAAGCCGGCTGTGCGGCAAGCTCATCCGCTA 496
Db 420 GGGCTTCGTGAAAGCAAGACGAGCAAGCCGGCTGTGCGGCAAGCTCATCCGCTA 479
Qy 497 CGCCAACTGGGCAACGTGCTCATCTGCGGACGCTCAGCACCGCAGTCTACAGCGCTT 556
Db 480 CGCCAACTGGGCAACGTGCTCATCTGCGGACGCTCAGCACCGCAGTCTACAGCGCTT 539
Qy 557 CCCAGGCGCCAGACCTGGTGGCAAGCAAGCTTTATGCTCCGGCAAGCAAGCAGTT 616
Db 540 CCCAGGCGCCAGACCTGGTGGCAAGCAAGCTTTATGCTCCGGCAAGCAAGCAGTT 599
Qy 617 GGAAGAACTGAGCTTACACACAATGTTCTGGGTGCTCGGTGTGTTGGCAACCT 676
Db 600 GGAAGAACTGAGCTTACACACAATGTTCTGGGTGCTCGGTGTGTTGGCAACCT 659
Qy 677 GTCAATGAAGCGTGGCTTGGAGGTGCAATCGGAACTCTATCTGCTCCAGAGCTGCT 736
Db 660 GTCAATGAAGCGTGGCTTGGAGGTGCAATCGGAACTCTATCTGCTCCAGAGCTGCT 719
Qy 737 GAACGAGATGAACAACCTTGGCTACTCACTGCTGAGCACTGTATGCTTACGATGATTA 796
Db 720 GAACGAGATGAACAACCTTGGCTACTCACTGCTGAGCACTGTATGCTTACGATGATTA 779
Qy 797 TATCCACTGCTGTATACACAGGTGTGACAGTGTGAGCACTGTATGCTTACGATGATTA 856
Db 780 TATCCACTGCTGTATACACAGGTGTGAGCACTGTATGCTTACGATGATTA 839
Qy 857 TCTAGTTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGTGCTTCTGACTTGG 916
Db 840 TCTAGTTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGTGCTTCTGACTTGG 899
Qy 917 CGTTGTGCGGCTCTTCAAGCTTCTGCAAGTTCTTCTATGTTGGCTGCTGAAGTGG 976
Db 900 CGTTGTGCGGCTCTTCAAGCTTCTGCAAGTTCTTCTATGTTGGCTGCTGAAGTGG 953
Qy 977 CCTCTCAGAGGCTGCTGAGCTGGAAGCATGGCAGAGGGGTATGGCCAGAGCTGCT 1036
Db 954 ----- 953
Qy 1037 TGAGACGAGATGCAAGTGTCAAGAAAGGAGTCTCACGGGTAGAAAGACGACGCGTG 1096
Db 954 ----- 953
Qy 1097 GTGGCGACACCTGTAATCCAGCTACTCGGGAAGCTGAGCAGAGAAATGCTTGAAC 1156
Db 954 ----- 953
Qy 1157 CGGAGGCGGAGGTTGTGGTGGCAGAGAGCTCATCAACCCCTTGGAGAGATGATGAT 1216
Db 954 -----GTTGGCAGAGAGCTCATCAACCCCTTGGAGAGATGATGAT 996
Qy 1217 GATTTTGAACCACTGATTTGTGCAAGAAATTTGACAGGATGTCCTGTGTTGGAT 1276
Db 997 GATTTTGAACCACTGATTTGTGCAAGAAATTTGACAGGATGTCCTGTGTTGGAT 1056
Qy 1277 GAGATGCAACGAGACCTGCTCGATGAGACCGGACATGTAAGTAATTAAGCCGAGCCA 1336
Db 1057 GAGATGCAACGAGACCTGCTCGATGAGACCGGACATGTAAGTAATTAAGCCGAGCCA 1116
Qy 1337 CAGCCGCCCTACACAGCTGCTCGGCCAAGTTCCGTGAGAGCTCTTATATGAGCTCCACC 1396
Db 1117 CAGCCGCCCTACACAGCTGCTCGGCCAAGTTCCGTGAGAGCTCTTATATGAGCTCCACC 1176
Qy 1397 TTCAACATCAAGCTTGAACAAAGAGATGAGTTTCAAGCCCAATCAGAGAGAGAGAG 1456
Db 1177 TTCAACATCAAGCTTGAACAAAGAGATGAGTTTCAAGCCCAATCAGAGAGAGAGAG 1236
Qy 1457 GATGCTCAAGCTGCAATTTGCGCTTCTTAGGCTGCAAGTCCCATGATCACAATCTCT 1516
Db 1237 GATGCTCAAGCTGCAATTTGCGCTTCTTAGGCTGCAAGTCCCATGATCACAATCTCT 1296
Qy 1517 CCCAGGCGAAACTGAAGACCAAACTAATGAGCCCAAGGGAATCCCTTCTCAAGAG 1576

|||||
Db 1297 CCGAGGGCAAACTCAAGACCAACTACTGTGGCCCAAGGGGATCCCTTCTCCAGSAG 1356
QY 1577 GGCGTCGCCAAAACCAAGGCGAGCCAAACAGAACTTTAGGGGCCAGAAACAAACAG 1636
Db 1357 GGCGTCGCCAAAACCAAGGCGAGCCAAACAGAACTTTAGGGGCCAGAAACAAACAG 1416
QY 1637 GGCGTGAAGCTTAAAGGCTGTGACGCGCTTCAAGTCTGGCCACCTGATTCAGAGGCCAGGC 1696
Db 1417 GGCGTGAAGCTTAAAGGCTGTGACGCGCTTCAAGTCTGGCCACCTGATTCAGAGGCCAGGC 1476
QY 1697 TACTACAGTGGCCCAACAGAGCGCCCTCAGGCCCACTGCCATGTTCTTCCCTAGAACCA 1756
Db 1477 TACTACAGTGGCCCAACAGAGCGCCCTCAGGCCCACTGCCATGTTCTTCCCTAGAACCA 1536
QY 1757 TCAGCGCCGTCGAAAGCTTCAAGTGTCAAGGATAGACACCAAAACAAAGCTTAAAG 1816
Db 1537 TCAGCGCCGTCGAAAGCTTCAAGTGTCAAGGATAGACACCAAAACAAAGCTTAAAG 1596
QY 1817 ACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTG 1876
Db 1597 ACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTG 1656
QY 1877 ATGAGACACCCAGAACTATCTCAAGTGAAGAGAAACCTGTGAGTTTAACTTGAACGAT 1936
Db 1657 ATGAGACACCCAGAACTATCTCAAGTGAAGAGAAACCTGTGAGTTTAACTTGAACGAT 1716
QY 1937 ATGCGAGAGATCCCGCAAAATCACTCAAGAAACCTTTGGAACAATCAACCAACATA 1996
Db 1717 ATGCGAGAGATCCCGCAAAATCACTCAAGAAACCTTTGGAACAATCAACCAACATA 1776
QY 1997 CACACACACTCAAAAGATCATCATGATCTTTATTTGGGCTTTGAAAAACAGGATGAACA 2056
Db 1777 CACACACACTCAAAAGATCATCATGATCTTTATTTGGGCTTTGAAAAACAGGATGAACA 1836
QY 2057 CATTCCTAAGCTGTCTCTAATGGGAGTGGCTTGGCAGCGAGTCTCACTGTGTGAC 2116
Db 1837 CATTCCTAAGCTGTCTCTAATGGGAGTGGCTTGGCAGCGAGTCTCACTGTGTGAC 1896
QY 2117 ACCAGAGAGACATGATCCAGTCAAGCAGCATACAGCTGTCCACTGAAAGACGTGTCT 2176
Db 1897 ACCAGAGAGACATGATCCAGTCAAGCAGCATACAGCTGTCCACTGAAAGACGTGTCT 1956
QY 2177 ACAACAGCGCTGAATCAAAATGTTAGCTTAATGATTAATAAATCCAGACTCTTCAAGCTT 2236
Db 1957 ACAACAGCGCTGAATCAAAATGTTAGCTTAATGATTAATAAATCCAGACTCTTCAAGCTT 2016
QY 2237 TAAATGCTTTTATTCATTAATAAATCTGTGAAGCTAGACTGAACATTTGAACT 2296
Db 2017 TAAATGCTTTTATTCATTAATAAATCTGTGAAGCTAGACTGAACATTTGAACT 2076
QY 2297 CAGACTCTGGAATTCAGAGTGGGAAACCTTTAGTTCTATCTGAATCCAAAGCAGCACACC 2356
Db 2077 CAGACTCTGGAATTCAGAGTGGGAAACCTTTAGTTCTATCTGAATCCAAAGCAGCACACC 2136
QY 2357 TTAGTATTAATGCGCAAACTAATAGTTTAATAA 2390
Db 2137 TTAGTATTAATGCGCAAACTAATAGTTTAATAA 2170

RESULT 7
AY357925 2187 bp mRNA linear PRI 26-AUG-2003
LOCUS Macaca fascicularis beestrophin (VMD2) mRNA, complete cds.
DEFINITION
ACCESSION AY357925
VERSION AY357925.1 GI:34013782
KEYWORDS
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE 1 (bases 1 to 2187)

AUTHORS Okamoto,H., Umeda,S., Suzuki,M.T., Yoshikawa,Y., Tanaka,Y. and Iwata,T.
TITLE Molecular Cloning of VMD2 Gene from Cynomolgus Monkey (Macaca fascicularis)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2187)
AUTHORS Okamoto,H., Umeda,S., Suzuki,M.T., Yoshikawa,Y., Tanaka,Y. and Iwata,T.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-2003) National Tokyo Medical Center, National Institute of Sensory Organs, 2-5-1, Higashiagoka, Meguro-ku, Tokyo 152-0021, Japan
FEATURES
source
location/Qualifiers
1..2187
/organism="Macaca fascicularis"
/mol_type="mRNA"
/db_xref="taxon:9541"
1..2187
/gene="VMD2"
67..1824
/gene="VMD2"
/codon_start=1
/product="beestrophin"
/protein_id="AA056049.1"
/db_xref="GI:34013783"
1..2187
/translation="MTITTSQYANARLGSFRRLLICWRGSIYKLLYGEFFIFLLCY IIRFIRLALTEOQLMEFKLLLYCDSYQLPISEVLGFVTLVVTWRMNDYENLPW PDRMLSVSGFVEGKDEQRLRLIRANIGNLILRSVSTAYAKFPPSHQILVOA GCMTPAEHKELEKLSLPHNMFVWPWPFNLKMKMLGSRIDPILLOSILNEMNTLR TQGHLYADWDIMLSPLVYQVTVAVYSFLLCLVGRQPLNPAKAYPGHEIDLIVPVF TPLQFFPVYGLMKVAEOLINPFGEDDDDEETWIDIRINOVSLIAVDEMDKPLPMEP DMYNPEPHPPTTAASQFRASFMGSPFNILNKEBEREPNOEDKDATIGILGR FIGLOSHDHPGANSRTKLIPKXBSLHLEBLPNHKAIVKONRGLSDNRMKIKAV DAKSPILYORPQIYSAPOTPLSPFPMFPPEPVSLSKHSVTGLIDTOKSLKTVSSG ANKSELSGSGALMEHPEVSHVRKVEFNLIDMPBILPENHLKEPELSTYNIHAT LNDHDPYALNBRDAHS"

ORIGIN
Query Match 74.5%; Score 1809.6; DB 9; Length 2187;
Best Local Similarity 87.5%; Pred. No. 0;
Matches 2092; Conservative 0; Mismatches 89; Indels 209; Gaps 2;
QY 39 GGATCATCGGACCACTGGAACCCACCTGACCCCAAGCCCACTGTGCAAGCCCACTGC 98
Db 1 GGATCACCGGATCCACTGGAACCCCACTGACCCCAAGCCCACTGTGCAAGCGTGC 60
QY 99 CTGGCCATGACCATCACTTACCAAGCCCAAGTGGCTAAATGCCGCTTAAGCTCTCTCC 158
Db 61 CTGGCCATGACCATCACTTACCAAGCCCAAGTGGCTAAATGCCGCTTAAGCTCTCTCC 120
QY 159 CGCTGTGCTGTGCTGCGGGGGCAGATCTAACAGCTGTATATGGGAGTTCCTTAATC 218
Db 121 CGCTGTGCTGTGCTGCGGGGGCAGATCTAACAGCTGTATATGGGAGTTCCTTAATC 180
QY 219 TTCTGTCTGTCTACTATCATCATCGCTTTATTTATAGCTGTGCGCTTACCGAAGAACAA 278
Db 181 TTCTGTCTGTCTACTATCATCATCGCTTTATTTATAGCTGTGCGCTTACCGAAGAACAA 240
QY 279 CAGCTGATGTTTGAAGAACTGATCTGTATTCGACAGCTATACATTCATCCCAATT 338
Db 241 CAGCTGATGTTTGAAGAACTGATCTGTATTCGACAGCTATACATTCATCCCAATT 300
QY 339 TCTTTCGTGTGGGCTTCTACGTACGCTGTGTGTGAGCCGCTGTGTGGAACCAAGTACGAG 398
Db 301 TCTTTCGTGTGGGCTTCTACGTACGCTGTGTGTGAGCCGCTGTGTGGAACCAAGTACGAG 360
QY 399 AACCTGCGGTGGCCCGACCGCTTATAGAGCTGTGTGTGAGGCTTGTGGAAGCAAGAC 458
Db 361 AACCTGCGGTGGCCCGACCGCTTATAGAGCTGTGTGTGAGGCTTGTGGAAGCAAGAC 420
QY 459 GAGCAAGGCGGCTGTGCGCGCGCAGCTCATCGCATCGCAAGCAACCTGGGCAACGCTGC 518
Db 421 GAGCAAGGCGGCTGTGCGCGCGCAGCTCATCGCATCGCAAGCAACCTGGGCAACGCTGC 480

QY 519 ATCTGCGCAGCGTCAGACCGCACTTACAGCGCTTCCCGAGGCCAGCACTGTG 578
DB 481 ATCTTGCGCAGCGTCAGACCGCGGTCTACAGCGCTTCCCGAGGCCAGCACTGTG 540
QY 579 CAAGCAGGCTTTATGACTCCGCGAGAACACAGCAGTTGAGAGAACTGAGCTTACAC 638
DB 541 CAACAGGCTTTATGACCCCGGAGAACACAGCAGTTGAGAGAACTGAGCTTACAC 600
QY 639 AACATGTTCTGGGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 698
DB 601 AACATGTTCTGGGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY 699 GGTGGAATCCGGGAGCCCTATCTGTCTCCAGAGCCGTGAGAGAGATGAACTGTGGT 758
DB 661 GGTGGAATCCGGGAGCCCTATCTGTCTCCAGAGCCGTGAGAGAGATGAACTGTGGT 720
QY 759 ACTGAGTGTGACACTGTGTATGCTGACAGCTGAGATTGATCCCACTGTGTATACAG 818
DB 721 ACTGAGTGTGACACTGTGTATGCTGACAGCTGAGATTGATCCCACTGTGTATACAG 780
QY 819 GTGTGTGACTGTGGCGGTGTACAGCTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGT 878
DB 781 GTGTGTGACTGTGGCGGTGTACAGCTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGT 840
QY 879 AACCCAGCCAAAGGCTTACCTGTGAGCTGAGCTGAGCTGTGTGTGTGTGTGTGTGTGT 938
DB 841 AACCCAGCCAAAGGCTTACCTGTGAGCTGAGCTGAGCTGTGTGTGTGTGTGTGTGTGT 900
QY 939 CTGCACTTCTTCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 998
DB 901 CTGCACTTCTTCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 932
QY 999 TGGAGGCAATGCGCAGAGGGGTCTATGCGCAGCAGCTGTGTGTGTGTGTGTGTGTGT 1058
DB 933 ----- 932
QY 1059 GAAAGGAAGGTCTACCGGGTAAAGACAGCCAGCGGTGTGTGTGTGTGTGTGTGTGT 1118
DB 933 ----- 932
QY 1119 GCTACTCGGAGGCTGAGCAGAGAAATCGCTTGAACCCGGAGGCGGAGTTGTGTGTGT 1178
DB 933 -----GTTGG 937
QY 1179 CAGAGCAGCTATCAACCCCTTTGAGAGAGATGATGATTTTGAAGCAACTGTGATTTG 1238
DB 938 CAGAGCAGCTATCAACCCCTTTGAGAGAGAGATGATGATTTTGAAGCAACTGTGATTTG 997
QY 1239 TCGACAGGAATTTGCAAGGTGTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1298
DB 998 TCGACAGGAATTTGCAAGGTGTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1057
QY 1299 GGATGAGCCGAGCACTGTACTGAAATTAAGCCGAGCCACAGCCCTTACACAGCTGTCTT 1358
DB 1058 GGATGAGCCGAGCACTGTACTGAAATTAAGCCGAGCCACAGCCCTTACACAGCTGTCTT 1117
QY 1359 CCGGCCAGTTCCGTGAGCTTCTTTATGGGCTCCACTTCAACATCACTGCTGAACAAG 1418
DB 1118 CCGGCCAGTTCCGTGAGCTTCTTTATGGGCTCCACTTCAACATCACTGCTGAACAAG 1177
QY 1419 AGGAGATGGAATTTCCAGCCCAATCAAGAGAGACGAGAGAGATGATGATGATGATGATG 1478
DB 1178 AGGAGATGGAATTTCCAGCCCAATCAAGAGAGACGAGAGAGATGATGATGATGATGATG 1237
QY 1479 GCGCGTTCTCTAGGCTGTGAGTCCCATGTATCAACATCTTCCAGAGGCAAACTCAAGAGCA 1538
DB 1238 GCGCGTTCTCTAGGCTGTGAGTCCCATGTATCAACATCTTCCAGAGGCAAACTCAAGAGCA 1297
QY 1539 AACTACTGTGGCCCAAGAGGAGATCTTCTTCCAGAGGCGCTGTCCCAAAAACCAAG 1598
DB 1298 AACTACTGTGGCCCAAGAGGAGATCTTCTTCCAGAGGCGCTGTCCCAAAAACCAAG 1357

QY 1599 CAGCCAAACAGACGTTAAGGGCCAGAGAGAACAAACAGGCGCTGAGAGCTTAAAGCTGTGTG 1658
DB 1358 CAGTCAAAACAGACGTTAAGGGCCAGAGAGAACAAACAGGCGCTGAGAGCTTAAAGCTGTAG 1417
QY 1659 ACGCTTCAAGTCTGAGCCCACTGTATCAAGAGCCAGGCTTACTACGTGCCCCACAGAGCC 1718
DB 1418 ATGCTTCAAGTCTGAGCCCACTGTATCAAGAGCCAGGCTTACTACGTGCCCCACAGAGCC 1477
QY 1719 CCGTCAGCCCACTGAGCTTCTTCCGCTTAAAGACATCAAGCGCGCTCAAAAGCTTTCACA 1778
DB 1478 CCGTCAGCCCACTGAGCTTCTTCCGCTTAAAGACATCAAGCGCGCTCAAAAGCTTTCACA 1537
QY 1779 GTGTCAACAGGATAGACACCAAGACCAAAAGCTTAAAGACTGTGAGTTCTGAGGCGCAAGA 1838
DB 1538 GTGTCAACAGGATAGACACCAAGACCAAAAGCTTAAAGACTGTGAGTTCTGAGGCGCAAGA 1597
QY 1839 AAAGTTTGAATTTGCTCTCAGAGAGCGATGAGGCGCTTATGAGAGCAACCAAGATATCTC 1898
DB 1598 AAAGTTTGAATTTGCTCTCAGAGAGCGATGAGGCGCTTATGAGAGCAACCAAGATATCTC 1657
QY 1899 AAGTGAAGGAGAAACGTGAGATTAACTGAGCGGATATGCGAGAGATCCCGGAAATTC 1958
DB 1658 AAGTGAAGGAGAAACGTGAGATTAACTGAGCGGATATGCGAGAGATCCCGGAAATTC 1717
QY 1959 AACTCAAAAGACCTTTGGAAACATCAACCAACATACACACTACCTCAAAAGATCA 2018
DB 1718 AACTCAAAAGACCTTTGGAAACATCAACCAACATACACACTACCTCAAAAGATCA 1777
QY 2019 TGGATCTTATTTGGGCTTGTGAAACACAGGATGAAGCAATTCCTTACCTGTCTTCTTAT 2078
DB 1778 TGGATCTTATTTGGGCTTGTGAAACACAGGATGAAGCAATTCCTTACCTGTCTTCTTAT 1837
QY 2079 GGGAGTGTTCGCGACGAGGCTTCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2138
DB 1838 GGGAGTGTTCGCGACGAGGCTTCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1897
QY 2139 CACAGCCATACAGCTGTCTCAGCTGAGAGAGATGATGATGATGATGATGATGATGATGATG 2198
DB 1898 CACAGCCATACAGCTGTCTCAGCTGAGAGAGATGATGATGATGATGATGATGATGATGATG 1957
QY 2199 TAGCTTAATGATTAATA-----AATCCAGACTACTTCAAGCTTTAATGCTTTTATTTCA 2252
DB 1958 TAGCTTAATGATTAATA-----AATCCAGACTACTTCAAGCTTTAATGCTTTTATTTCA 2017
QY 2253 TAAAAACGTGAAGCTGTGATGAAACATTTGAACATTTAATCTGAGCTGTGATTTGAG 2312
DB 2018 TAAAAATTTATGAAGCTGTGATGAAACATTTGAACATTTAATCTGAGCTGTGATTTGAG 2077
QY 2313 AGTGGGGAACCTTATGTTCTATCTGATCAATCAAGACGACCACTTATATCTGCCAA 2372
DB 2078 AGTGGGGAACCTTATGTTCTATCTGATCAATCAAGACGACCACTTATATCTGCCAA 2137
QY 2373 ACTAATGATTTAATTAATTAACAAATCTGTTAAAAAAGGAGGAGGAGGAGGAGGAGGAGG 2422
DB 2138 ACTAATGATTTAATTAATTAACAAATCTGTTAAAAAAGGAGGAGGAGGAGGAGGAGGAGG 2187

RESULT 8
BC015220 2441 bp mRNA linear PRI 07-OCT-2003
LOCUS Homo sapiens vitelliniform macular dystrophy (Best disease,
bestrophin), mRNA (cdna clone MGC:117761 IMAGE:3877806), complete
cde.
ACCESSION BC015220
VERSION BC015220.2 GI:34783515
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2441)
Straubeberg,R.L., Feingold,E.A., Grouse,L.H., Dergs,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenneman,C.M., Schuler,G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stedman, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ueda, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., Wierley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E., Schmechel, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

JOURNAL MEDLINE PubMed 12477932

2 (bases 1 to 2441)

Strausberg, R.

Direct Submission

Submitted (01-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Sep 16, 2003 this sequence version replaced gi:21955361.

Contact: MGC help desk

Email: cgabs-remail.nih.gov

Tissue Procurement: DCTD/DTF/Gazdar

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www.shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>

Series: IRAX Plate: 14 Row: b Column: 9.

location/Qualifiers

1. 2441

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:17761 IMAGE:3877806"

/tissue_type="lung, large cell carcinoma"

/clone_lib="NIH_MGC_68"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

1. 2441

/gene="VMD2"

/note="synonyms: BMD, BRST"

/db_xref="locusid:7439"

/db_xref="MIM:153700"

573. 2069

/gene="VMD2"

/codon_start=1

/product="VMD2 protein"

/protein_id="AAH15220.1"

/db_xref="GI:22800661"

/db_xref="Locusid:7439"

/db_xref="MIM:153700"

/translation="MEKTLTYCDSYTIPLISFVIGFYVTLVTPMNOYNTLPMWPRLMSIVSGFVEKDEGRILRLIRYANLGNVLIRSYSTAYRFRPAQHIVQAGFMTPAEHKOLEKSLPHNMFVPPVWPAISKAMLGKRLDPLILDSLINTENNTLRQCGHLYADWISTPLVTVQVTVAVISFPLTCLVGRQFLNPARAPGHEIDLVPVPVTFPLQFFFYVGMILKSLVLAIDEMHODLPMEEDPMWNKDEPPPYTAASAGFRASFMGST

misc_feature

576. 1427

/gene="VMD2"

/note="DUF289, Region: Putative membrane protein. This family called family 8 in, may be a transmembrane protein The specific function of this protein is unknown"

/db_xref="CDD:pfam01062"

ORIGIN

Query Match 65.3%; Score 1587; DB 9; Length 2441;

Beet Local Similarity 86.7%; Pred. No. 0;

Matches 1884; Conservative 0; Mismatches 5; Indels 284; Gaps 1;

257 GCTGGCCCTCAGGAGAAACAAGCTGATGTTTGAAGAACTGCTGATTTGGCAG 316

545 GCTGGCCCTCAGGAGAAACAAGCTGATGTTTGAAGAACTGCTGATTTGGCAG 604

317 CTACATCCAGCTCATCCCAATTTCTTGTGCTGGGCTTCTAGCTGACGCTGCTGAC 376

605 CTACATCCAGCTCATCCCAATTTCTTGTGCTGGGCTTCTAGCTGACGCTGCTGAC 664

377 CCGCTGGTGAACACGATGAGAAACCTGCGCGTGGCCGACCGCTTATGAGCTGTGTC 436

665 CCGCTGGTGAACACGATGAGAAACCTGCGCGTGGCCGACCGCTTATGAGCTGTGTC 724

437 GGGCTTGTGGAAGCAAGGACGAGCAAGGCGCGCTGTGCGGCGACGCTCATCCGCTA 496

725 GGGCTTGTGGAAGCAAGGACGAGCAAGGCGCGCTGTGCGGCGACGCTCATCCGCTA 784

497 CGCAACCTGGGGAAGAGTGTCTCTCTGGCAGCTTACAGACGCTTACAGCGCTT 556

785 CGCAACCTGGGGAAGAGTGTCTCTCTGGCAGCTTACAGACGCTTACAGCGCTT 844

557 CCCAGGCGCCAGCACTGCTGAGCAAGGCTTATGATCTCGGCGAGAACAGAGAGT 616

845 CCCAGGCGCCAGCACTGCTGAGCAAGGCTTATGATCTCGGCGAGAACAGAGAGT 904

617 GGAAGAACTGAGCTTACCAACAACATGTTGCGGCTGCGGCTGTTGCCAAGCT 676

905 GGAAGAACTGAGCTTACCAACAACATGTTGCGGCTGCGGCTGTTGCCAAGCT 964

677 GTCAATGAAGGCGTGTGAGGTGCAATCCGGGACCTTACCTGCTCCAGAGCTGCT 736

965 GTCAATGAAGGCGTGTGAGGTGCAATCCGGGACCTTACCTGCTCCAGAGCTGCT 1024

737 GAAAGATGAACACCTTGCCTGACTCAGTGTGACACCTGTATGCTACGACTGATTAG 796

1025 GAAAGATGAACACCTTGCCTGACTCAGTGTGACACCTGTATGCTACGACTGATTAG 1084

797 TATCCACAGTGATATACAGAGTGTGATCTGTGCGGTGTACAGCTTCTCTGACTTG 856

1085 TATCCACAGTGATATACAGAGTGTGATCTGTGCGGTGTACAGCTTCTCTGACTTG 1144

857 TCTAGTTGGCGCGAGTTTCTGAACCCAGCAAGGCTTACCTGCGGCTGAGCTGACT 916

1145 TCTAGTTGGCGCGAGTTTCTGAACCCAGCAAGGCTTACCTGCGGCTGAGCTGACT 1204

917 CGTTGTGCGCGCTTTCACGCTTCTGCACTTCTTCTTATGTTGGCTGAGAGTGG 976

1205 CGTTGTGCGCGCTTTCACGCTTCTGCACTTCTTCTTATGTTGGCTGAGAGTGG 1257

977 CCTTCCAGGGCCCTGCTGGGCTGGAAGCAATGGGCAAGGGGTATGAGCCAGAGCTGCT 1036

1258 ----- 1257

1037 TGAGACGAGATGACGTGTCAAGAAAGAAAGTCTACGCGGTAGAAAGACGAGCGGTG 1096

1258 ----- 1257

1097 GTGGCGACACCTTAATCCAGCTACTCGGAGAGCTGAGGACGAGAGATGCTTGAACC 1156

```

Db      1258 ----- 1257
Qy      1157 CGGAGGCGGAGTGTGTGTGGCAGAGAGCTCATCAACCCCTTTGGAGAGATGATGAT 1216
Db      1258 ----- 1257
Qy      1217 GATTTTGAGACCAACTGAGATTGTGCAGAGAAATTTGACAGTGTCCCTGTGTGGCTGTGAT 1276
Db      1258 ----- 1280
Qy      1277 GAGATGACACAGAGACTGTCTCGATGAGAGCCCGACATGTACTGTGAATPAAGCCGAGCCA 1336
Db      1281 GAGATGACACAGAGACTGTCTCGATGAGAGCCCGACATGTACTGTGAATPAAGCCGAGCCA 1340
Qy      1337 CAGGCCCCCTTACAGAGCTGTCCGCGCCAGTTCCTGAGGCTCCTTTAATGGGCTCCACC 1396
Db      1341 CAGCCCCCTTACAGAGCTGTCCGCGCCAGTTCCTGAGGCTCCTTTAATGGGCTCCACC 1400
Qy      1397 TTCAACATCAGCCTGAAACAAAGAGAGATGAGATTCCAGCCCAATCAGAGAGAGAG 1456
Db      1401 TTCAACATCAGCCTGAAACAAAGAGAGATGAGATTCCAGCCCAATCAGAGAGAGAGAG 1460
Qy      1457 GATGCTCAGCGTGGCATGATGGCCGCTTCTTAAGGCTGAGTCCATGATCACCATCT 1516
Db      1461 GATGCTCAGCGTGGCATGATGGCCGCTTCTTAAGGCTGAGTCCATGATCACCATCT 1520
Qy      1517 CCCAGGCGCAACTCAAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTTCTCCACAG 1576
Db      1521 CCCAGGCGCAACTCAAGACCAAACTACTGTGGCCCAAGAGAGAAATCCCTTCTCCACAG 1580
Qy      1577 GGCCTGCGCCAAACCAACAGGACGACCAAGAACGATTAGGGGCGAGAGAACAAACAG 1636
Db      1581 GGCCTGCGCCAAACCAACAGGACGACCAAGAACGATTAGGGGCGAGAGAACAAACAG 1640
Qy      1637 GCCTGGAAGCTTAAAGGCTGTGAGACGCTTCAAGTCTGCGCCCACTGTATCAGAGGCGAGC 1696
Db      1641 GCCTGGAAGCTTAAAGGCTGTGAGACGCTTCAAGTCTGCGCCCACTGTATCAGAGGCGAGC 1700
Qy      1697 TACTACAGTGGCCCAACAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCTTGAACCA 1756
Db      1701 TACTACAGTGGCCCAACAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCTTGAACCA 1760
Qy      1757 TCAGGCGCGTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAGAACAAAGCTTAAAG 1816
Db      1761 TCAGGCGCGTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAGAACAAAGCTTAAAG 1820
Qy      1817 ACTGTGAGTCTGGGCGCCAAAGAAATTTGAATTTGCTCTCAGAGAGCGATGGGCGCTTG 1876
Db      1821 ACTGTGAGTCTGGGCGCCAAAGAAATTTGAATTTGCTCTCAGAGAGCGATGGGCGCTTG 1880
Qy      1877 ATGAGACCCCAAGAAATTTCTCAAGTGAAGAGAGAAATCTGTGAGATTAACTCTGACGAT 1936
Db      1881 ATGAGACCCCAAGAAATTTCTCAAGTGAAGAGAGAAATCTGTGAGATTAACTCTGACGAT 1940
Qy      1937 ATGCGAGAGATCCCCGAAATACCTCTCAAGAACCTTTGGAACATGACCAACCAACTA 1996
Db      1941 ATGCGAGAGATCCCCGAAATACCTCTCAAGAACCTTTGGAACATGACCAACCAACTA 2000
Qy      1997 CACACTACACTCAAAAGATCAGATGATCTTATTGGGCTTTGAAAAAGGAGATGAAGCA 2056
Db      2001 CACACTACACTCAAAAGATCAGATGATCTTATTGGGCTTTGAAAAAGGAGATGAAGCA 2060
Qy      2057 CATTCTCAACTGCTTCTTAATGGGAGTCTTCCGACGCGAGTCTCACTGTGTGTAC 2116
Db      2061 CATTCTCAACTGCTTCTTAATGGGAGTCTTCCGACGCGAGTCTCACTGTGTGTAC 2120
Qy      2117 ACCAGCAGGACACTGATCAGTCAAGCATTACAGCTGTGACACTGAAGAAAGTGTCT 2176
Db      2121 ACCAGCAGGACACTGATCAGTCAAGCATTACAGCTGTGACACTGAAGAAAGTGTCT 2180
Qy      2177 ACAACAGCTGAATCAAAATGTTAGCTTAATGATTAATAATCCGAGACTACCTTGACCTT 2236

```

```

Db      2181 ACAACAGCTGAATCAAAATGTTAGCTTAATGATTAATAATCCAGACTACTGACCTT 2240
Qy      2237 TAATGCCCTTTATTCATTAATAAATCTGTGAAGACTGAGTGAACCAATGGAACATTAAT 2296
Db      2241 TAATGCCCTTTATTCATTAATAAATCTGTGAAGACTGAGTGAACCAATGGAACATTAAT 2300
Qy      2297 CAGACTCTGATTCAGAGTGGGAAACCTTATGTTCTATCTGAATTCAGAGCGCACACC 2356
Db      2301 CAGACTCTGATTCAGAGTGGGAAACCTTATGTTCTATCTGAATTCAGAGCGCACACC 2360
Qy      2357 TTAGTATATGCGCCCAACTAATGAGTTTAATTAATTAACAAATACCTGTTAAAAA 2416
Db      2361 TTAGTATATGCGCCCAACTAATGAGTTTAATTAATTAACAAATACCTGTTAAAAA 2420
Qy      2417 AAAAAAAAAA 2429
Db      2421 AAAAAAAAAA 2433

```

```

RESULT 9
LOCUS   AY515704
DEFINITION Homo sapiens beestrophin 1 mRNA, complete cds.
VERSION AY515704.1
KEYWORDS
SOURCE  Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Sun,H., Tsunenari,T., Yau,K.W., and Nathans,J.
TITLE   The vitelliform macular dystrophy protein defines a new family of chloride channels
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (6), 4008-4013 (2002)
MEDLINE 21902086
PUBMED 11904445
REFERENCE
AUTHORS Yau,K.W. and Nathans,J.
TITLE   Structure-function analysis of the beestrophin family of anion channels
JOURNAL J. Biol. Chem. 278 (42), 41114-41125 (2003)
MEDLINE 22917481
PUBMED 12907679
REFERENCE
AUTHORS Sun,H., Tsunenari,T., Yau,K.-W. and Nathans,J.
TITLE   Direct Submission
JOURNAL Submitted (02-JAN-2004) Molecular Biology, Neuroscience, Johns Hopkins University, 725 N. Wolfe Street, Baltimore, MD 21205, USA
FEATURES
Source
1..1758
/mol_type="Homo sapiens"
/db_xref="taxon:9606"
1..1758
/note="chloride channel"
/codon_start=1
/product="beestrophin 1"
/protein_id="AA89654.1"
/db_xref="GI:41216873"
/translacion="MTITYSOVNANRLGFSRLLCMRGSIYKLYLGEFLLPILCY
IIRIVRIALTEBOQLMPEKTLTLCDSYICLIPSPFAGPVYTLVTRMNOYENLPM
PDRMLSVSGFVSGKDEOGRILRTTLIRYALGNVLLRVSIVTVYKRFPSAQHVOA
GFMTPAEHKQLEKLSLPHNMFVWVFWFANLSMAMWLGRIKDLILQSLNENWTLR
TOCHILAYADMIISILVYTQVTVAVVSFETLCVGFQPLPAAYAGHEDLIVVVF
TFLOPFYVGMILKVAEQLINFGEDDDDFENMIVDNLQVSLAVBHQDLPRMBP
DMWNKREPPRPYTAASOFRPASFMGSEFTNLSLXKEMEQPOBDEBDNAHAGIGR
FLGLOSHDHPRPASNSTKLIMRGESLIRHGLPKHKAAKONVREGQEDNKAMKLVK
DAKFSAPLYGRPYSHAPQTPLETPMFVFLBPSAPSKHSLVTCIDTKDSLKYVSSG
AKKSFELSSDGLMEHPVSOVRRRTVEFNLTDWELIPENHLKEPLEOSPTNIHTT
LKDHMDPYMALENRDEAHS"
ORIGIN

```

Query Match 63.5%; Score 1543.4; DB 9; Length 1758;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 1757; Conservative 0; Mismatches 1; Indels 203; Gaps 1;

```
QY 105 ATGACCATCTTACACAAAGCAAGGCTAATGCCGCTTAAAGCTCTTCCCGCTG 164
Db 1 ATGACCATCTTACACAAAGCAAGGCTAATGCCGCTTAAAGCTCTTCCCGCTG 60
QY 165 CTGCTGCTGCTGGGGGAGAGATCTACAGCTGTATATAGCCAGTTCTTAATCTTCG 224
Db 61 CTGCTGCTGCTGGGGGAGAGATCTACAGCTGTATATAGCCAGTTCTTAATCTTCG 120
QY 225 CTGCTACTACATCATCCGCTTATTTATATAGCTGCGCTCAAGAAAGCAAGCTG 284
Db 121 CTGCTACTACATCATCCGCTTATTTATATAGCTGCGCTCAAGAAAGCAAGCTG 180
QY 285 ATGTTTGAAGAACTGATCTGTATTTGGGACAGCTACATCCATTTCTTC 344
Db 181 ATGTTTGAAGAACTGATCTGTATTTGGGACAGCTACATCCATTTCTTC 240
QY 345 GTGCTGGCTTCTACGTGACGTGTGTGACCCGCGGTGGAAACAGTACAGAACTTG 404
Db 241 GTGCTGGCTTCTACGTGACGTGTGTGACCCGCTGGTGGAAACAGTACAGAACTTG 300
QY 405 CCGTGGCCGACCGCTCATGAGCTGTGTGCGGCTTGTGCGAAGGCAAGGACGAGCA 464
Db 301 CCGTGGCCGACCGCTCATGAGCTGTGTGCGGCTTGTGCGAAGGCAAGGACGAGCA 360
QY 465 GGCCTGCTGCTGGGCGGACGCTCATCCGCTACGCCAACCTGGGCAAGCTGTCTCTG 524
Db 361 GGCCTGCTGCTGGGCGGACGCTCATCCGCTACGCCAACCTGGGCAAGCTGTCTCTG 420
QY 525 CCGAGCGTGAACAACGCACTTACAGCGCTTCCCGAGCGCCAGCACTGGTGAAGCA 584
Db 421 CCGAGCGTGAACAACGCACTTACAGCGCTTCCCGAGCGCCAGCACTGGTGAAGCA 480
QY 585 GGGTTATATATCCCGGCAAGCAAGGCTTGGAGAACTGAGCTTACCAACAACATG 644
Db 481 GGGTTATATATCCCGGCAAGCAAGGCTTGGAGAACTGAGCTTACCAACAACATG 540
QY 645 TTCTGGGTGCTGAGGTGTGTGTTGCCAACCTGTCAATGAAGAGGTGGAGGTGCA 704
Db 541 TTCTGGGTGCTGAGGTGTGTGTTGCCAACCTGTCAATGAAGAGGTGGAGGTGCA 600
QY 705 ATCCGGAACCTTATCTCTGTCCAGAGCTGTGAACGAGATGAACCTTGCATCTAG 764
Db 601 ATCCGGAACCTTATCTCTGTCCAGAGCTGTGAACGAGATGAACCTTGCATCTAG 660
QY 765 TGTGGAACACTGTATGCTTACGATGATTTAGTATCCCATGTGTATATACACAGTGTG 824
Db 661 TGTGGAACACTGTATGCTTACGATGATTTAGTATCCCATGTGTATATACACAGTGTG 720
QY 825 ACTGTGGGTGTACAGCTTCTTCTGACCTTGTCTAGTTGGGCGGAGTTTCTGACCA 884
Db 721 ACTGTGGGTGTACAGCTTCTTCTGACCTTGTCTAGTTGGGCGGAGTTTCTGACCA 780
QY 885 GCAAGAGCTTACCTGCGGCATAGACTGAACCTGTGTGCGCGCTTCAAGTCTTGAG 944
Db 781 GCAAGAGCTTACCTGCGGCATAGACTGAACCTGTGTGCGCGCTTCAAGTCTTGAG 840
QY 945 TTCTTTCTTATATGTTGGCTGAGAGTGGGCTTCTCAGGGGCTGTGGGCTGGAAG 1004
Db 841 TTCTTTCTTATATGTTGGCTGAGAGTGGGCTTCTCAGGGGCTGTGGGCTGGAAG 866
QY 1005 CATGGCCAGAGGGGTCAATGGCCAGACAGCTGTGAGACGAGATGCAAGTGCAGAAAG 1064
Db 867 ----- 866
QY 1065 AAGGCTTCAAGGATGAAGAGCAGAGCGGTGTGGCGCACACCTGTAAATCCAGCTACT 1124
Db 867 ----- 866
```

```
QY 1125 CCGGAGGCTGAGGAGAGAAATCGCTTGAAACCCGAGGCGAGAGTTGTGGTGCAGAGC 1184
Db 867 -----GGTGCAGAGC 877
QY 1185 AGCTATGAACCCCTTGGAGAGATGATGATTTTGAAGCAACTGATTTGCA 1244
Db 878 AGCTATGAACCCCTTGGAGAGATGATGATTTTGAAGCAACTGATTTGCA 937
QY 1245 GGAATTTGACAGATGCTCTGTGTGGCTGTGATGATGACACAGAGCTGCTGGATG 1304
Db 938 GGAATTTGACAGATGCTCTGTGTGGCTGTGATGATGACACAGAGCTGCTGGATG 997
QY 1305 AGCCGAGATATGATCTGGAATTAAGCCGAGGCCAGCCCCCTTACACAGCTGTCCGCC 1364
Db 998 AGCCGAGATATGATCTGGAATTAAGCCGAGGCCAGCCCCCTTACACAGCTGTCCGCC 1057
QY 1365 AGTTCCGTCGAGCTCTCTTATGAGGCTCACTTCAACATTCAGCTGGAAGAAAGAGA 1424
Db 1058 AGTTCCGTCGAGCTCTCTTATGAGGCTCACTTCAACATTCAGCTGGAAGAAAGAGA 1117
QY 1425 TGAAGTTCCAGCCCAATCAGAGAGCAGAGAGATGCTCAGCTGCAATATGGCGCT 1484
Db 1118 TGAAGTTCCAGCCCAATCAGAGAGCAGAGAGATGCTCAGCTGCAATATGGCGCT 1177
QY 1485 TCCTAGGCTGCAAGTCCCATGATCACTATCTTCCAGGGCAACTCAAGGACCAACTAC 1544
Db 1178 TCCTAGGCTGCAAGTCCCATGATCACTATCTTCCAGGGCAACTCAAGGACCAACTAC 1237
QY 1545 TGTGGCCCAAGAGGGAATCCCTTCTCCAGAGGGGCTGCCAAAACCAAGAGCGCA 1604
Db 1238 TGTGGCCCAAGAGGGAATCCCTTCTCCAGAGGGGCTGCCAAAACCAAGAGCGCA 1297
QY 1605 AACGAAACCTTAAAGGCGCCAGAAAGCAACAAGGCTTGAAGCTTGAAGCGCT 1664
Db 1298 AACGAAACCTTAAAGGCGCCAGAAAGCAACAAGGCTTGAAGCTTGAAGCGCT 1357
QY 1665 TCAAGTCTGCGCCCACTGTATCAGAGGCGAGGCTATCAAGTCCCAAGAGCGCCCTCA 1724
Db 1358 TCAAGTCTGCGCCCACTGTATCAGAGGCGAGGCTATCAAGTCCCAAGAGCGCCCTCA 1417
QY 1725 GCGCCACTCCCATGTTCTTCCCTTGAACCATCAGCGCGCTCAAGCTTCAAGTGTCA 1784
Db 1418 GCGCCACTCCCATGTTCTTCCCTTGAACCATCAGCGCGCTCAAGCTTCAAGTGTCA 1477
QY 1785 CAGGCAATGACACCAAGCAAAAGCTTAAAGCTGTAGATTTGGGCGCCAAAGATT 1844
Db 1478 CAGGCAATGACACCAAGCAAAAGCTTAAAGCTGTAGATTTGGGCGCCAAAGATT 1537
QY 1845 TTGAATTGCTCTCAGAGAGCAGATGGGGCTTGATGAGACACCAAGATATCTCAAGTGA 1904
Db 1538 TTGAATTGCTCTCAGAGAGCAGATGGGGCTTGATGAGACACCAAGATATCTCAAGTGA 1597
QY 1905 GGAAGAAAACCTGTGAGATTTAACTGACGATATGCGAGATGCCGAATCCGAAATCACTCA 1964
Db 1598 GGAAGAAAACCTGTGAGATTTAACTGACGATATGCGAGATGCCGAATCCGAAATCACTCA 1657
QY 1965 AAGAACCTTTGGGAATCAACCAACCAATACATACATCACTCAAGATACATGATC 2024
Db 1658 AAGAACCTTTGGGAATCAACCAACCAATACATACATCACTCAAGATACATGATC 1717
QY 2025 CTTATTTGGGCTTGGAAAAAGGAGATGAAGCAATTCCTAA 2065
Db 1718 CTTATTTGGGCTTGGAAAAAGGAGATGAAGCAATTCCTAA 1758
```

RESULT 10
AX745964
LOCUS AX745964 1758 bp DNA linear PAT 14-MAY-2003
DEFINITION Sequence 1 from Patent WO03030922.
ACCESSION AX745964
VERSION AX745964.1 GI:30724619
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Stuenkel, A., Bruenner, G., Fritsch, R., Eulenberger, K. and
Ciosek, I.
TITLE Beetrophin and beetrophin homologous proteins involved in the
regulation of energy homeostasis
JOURNAL Patent: WO 03030922-A 1 17-Apr-2003;
Deutscher Patentschrift fuer Entwicklungsbiologische Forschung
(DE)
FEATURES
Source Location/Qualifiers
1..1758
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 63.5%; Score 1541.8; DB 6; Length 1758;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 1756; Conservative 0; Mismatches 2; Indels 203; Gaps 1;
QY 105 ATGACATCACTTACACAGCAAGCAAGTGGCTATGCGCGCTTACGCTTCTCCGCTG 164
Db 1 ATGACATCACTTACACAGCAAGCAAGTGGCTATGCGCGCTTACGCTTCTCCGCTG 60
QY 165 CTGCTGTCGCGCGGAGCAGCATCTACAGAGCTGTATAGGAGAGTTCTTATCTTCTG 224
Db 61 CTGCTGTCGCGCGGAGCAGCATCTACAGAGCTGTATAGGAGAGTTCTTATCTTCTG 120
QY 225 CTGCTGCTACATCATCCGCTTATTTATAGGCTGCGCTTACAGAGCAAGAGCTG 284
Db 121 CTGCTGCTACATCATCCGCTTATTTATAGGCTGCGCTTACAGAGCAAGAGCTG 180
QY 285 ATGTTTGAGAACTGACTCTGTATTTGGAGAGCTACATCCAGCTATCCCATTTCTTC 344
Db 181 ATGTTTGAGAACTGACTCTGTATTTGGAGAGCTACATCCAGCTATCCCATTTCTTC 240
QY 345 GTGCTGGAGCTTCTAGTACGCTGTGTGACCCGCTGTGGAAACAGTAGAGAACTG 404
Db 241 GTGCTGGAGCTTCTAGTACGCTGTGTGACCCGCTGTGGAAACAGTAGAGAACTG 300
QY 405 CCGTGGCCCGACCGCTCATAGAGCTGTGTGCGGCTTCTGAGAGGAGAGAGAGCA 464
Db 301 CCGTGGCCCGACCGCTCATAGAGCTGTGTGCGGCTTCTGAGAGGAGAGAGAGCA 360
QY 465 GCGCGGCTGCTGCGGAGCAAGCTCATCCGCTACGCAACTGGGCACTGTCTATCTG 524
Db 361 GCGCGGCTGCTGCGGAGCAAGCTCATCCGCTACGCAACTGGGCACTGTCTATCTG 420
QY 525 CGAGAGCTCAGACCGAGCTTACAGAGGCTTCCCGAGGCGCCAGCACTGTGTGCAACA 584
Db 421 CGAGAGCTCAGACCGAGCTTACAGAGGCTTCCCGAGGCGCCAGCACTGTGTGCAACA 480
QY 585 GCGTTTATGACTCCGAGCAAGCAAGAGCTTGGAGAACTGAGCTTACACAGCACTG 644
Db 481 GCGTTTATGACTCCGAGCAAGCAAGAGCTTGGAGAACTGAGCTTACACAGCACTG 540
QY 645 TTCTGGGCTGCTGCGGCTGTGCTTCCCACTGTCAATGAAGCGTGTGAGAGTGA 704
Db 541 TTCTGGGCTGCTGCGGCTGTGCTTCCCACTGTCAATGAAGCGTGTGAGAGTGA 600
QY 705 ATCGGGAGCCCTATCCGCTCAGAGAGCTGTGAAGAGAGAGAGAGAGAGAGAGAG 764
Db 601 ATCGGGAGCCCTATCCGCTCAGAGAGCTGTGAAGAGAGAGAGAGAGAGAGAGAG 660
QY 765 TGTGAGACACTGTATGCTTACAGAGCTGTATGATATCCCACTGTGTATACAGAGTGTG 824
Db 661 TGTGAGACACTGTATGCTTACAGAGCTGTATGATATCCCACTGTGTATACAGAGTGTG 720
QY 825 ACTGTGGGCTGTACAGCTTCTTCTGACTGTGTGAGTTGGGCGGAGTTTCTGAGCCA 884
Db 721 ACTGTGGGCTGTACAGCTTCTTCTGACTGTGTGAGTTGGGCGGAGTTTCTGAGCCA 780

QY 885 GCAAGGCTTACCTTGGCCATAGAGCTGAGCTGTGTGCGCGCTTCTTACAGTTCTTGGAG 944
Db 781 GCAAGGCTTACCTTGGCCATAGAGCTGAGCTGTGTGCGCGCTTCTTACAGTTCTTGGAG 840
QY 945 TTCTTCTTATGTTGGCTGTGAGAGTGGAGCTTCTTCAAGGCGCTGTGGGCTTGGAG 1004
Db 841 TTCTTCTTATGTTGGCTGTGAGAGTGGAGCTTCTTCAAGGCGCTGTGGGCTTGGAG 866
QY 1005 CATGGCCAGAGGGGCTATGCGCAGAGCTGCTTGAAGAGAGATGACGTTCAGAGAAAG 1064
Db 867 ----- 866
QY 1065 AAGGTCTACGGGTAGAAAGAGAGCCAGGCTGTGTGGCCAGACACTGATCCAGCTACT 1124
Db 867 ----- 866
QY 1125 CGGAGGCTGAGAGAGAGAAATCGCTTGAACCCCGGAGCGGAGGTTGTGTGAGAGC 1184
Db 867 -----GGTGTGAGAGC 877
QY 1185 AGCTCATCAACCCCTTGGAGAGATGATGATTTTGAACCACTGATTTGTGACA 1244
Db 878 AGCTCATCAACCCCTTGGAGAGATGATGATTTTGAACCACTGATTTGTGACA 937
QY 1245 GGAATTTGCAAGTGTCCCTGTGTGTGATGATGATGATGATGATGATGATGATGATG 1304
Db 938 GGAATTTGCAAGTGTCCCTGTGTGTGATGATGATGATGATGATGATGATGATGATG 997
QY 1305 AGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1364
Db 998 AGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1057
QY 1365 AGTTCCGTCGAGCTTCTTATGAGCTTCACTTCAATCAATCAATCAATCAATCAATCA 1424
Db 1058 AGTTCCGTCGAGCTTCTTATGAGCTTCACTTCAATCAATCAATCAATCAATCAATCA 1117
QY 1425 TGAAGTTCAAGCCCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1484
Db 1118 TGAAGTTCAAGCCCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1177
QY 1485 TCTTAGGCTTCAAGTCCCATGATCACTTCTTCCAGAGGCAATCAAGAGCAATCAAT 1544
Db 1178 TCTTAGGCTTCAAGTCCCATGATCACTTCTTCCAGAGGCAATCAAGAGCAATCAAT 1237
QY 1545 TGTGGCCCAAGAGGAATCCCTTCTTCCAGAGGCTTGTCCCAAAACCAAGAGAGCA 1604
Db 1238 TGTGGCCCAAGAGGAATCCCTTCTTCCAGAGGCTTGTCCCAAAACCAAGAGAGCA 1297
QY 1605 AACGAACTTAAAGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1664
Db 1298 AACGAACTTAAAGGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
QY 1665 TCAAGTCTGGCCCACTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1724
Db 1358 TCAAGTCTGGCCCACTGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1417
QY 1725 GCGCCACTCCCATGTTCTTCCCTTGAACCATCAGAGGCGGTCAAGTTCAAGTGTCA 1784
Db 1418 GCGCCACTCCCATGTTCTTCCCTTGAACCATCAGAGGCGGTCAAGTGTCA 1477
QY 1785 CAGGATTAAGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1844
Db 1478 CAGGATTAAGACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1537
QY 1845 TTGAATGCTTCTCAG 1904
Db 1538 TTGAATGCTTCTCAG 1597
QY 1905 GAG 1964
Db 1598 GAG 1657

QY	1655	AGAACCTTTGGAAACATACCAACCAACATACACATACATCAAGTCAATGATGC	2024
Db	1658	AGGAACCTTTGGAAACATACCAACCAACATACACATCAATCAAGATCATGATC	1717
QY	2025	CTTATTTGGGCTTTGGAAAAACAGGATGGAACATCTCTTAA	2065
Db	1718	CTTATTTGGGCTTTGGAAAAACAGGATGGAACATCTCTTAA	1758
RESULT 11			
LOCUS	BC041664		
DEFINITION	BC041664	2435 bp	mRNA linear
			PRI 07-OCT-2003
			Homio sapiens vitelliform macular dystrophy (Best disease, bestrorophin), mRNA (CDNA clone MGC:47884 IMAGE:5194643), complete cds.
ACCESSION	BC041664		
VERSION	BC041664.1	GI:27371319	
KEYWORDS	MGC.		
SOURCE	Homio sapiens (human)		
ORGANISM	Homio sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
REFERENCE	1 (bases 1 to 2435)		
AUTHORS	Strausberg, R. L., Feingold, E. A., Grouse, L. H., Derge, J. G., Klausner, R. D., Collins, F. S., Wagner, L., Shenmen, C. M., Schlier, G. D., Aleschul, S. F., Zeeberg, B., Blotow, K. H., Shaffer, C. F., Bhat, N. K., Hopkins, R. F., Jordan, H., Moore, T., Max, S. I., Wang, J., Hsieh, F., Dietchenko, L., Marusina, K., Farmer, A. A., Rubin, G. M., Hong, L., Stapleton, M., Soares, M. B., Donald, M. F., Casavant, T. L., Scheetz, T. E., Brownstein, M. J., Ushid, T. B., Toshiyuki, S., Carinelli, P., Prange, C., Raha, S. S., Loquellano, N. A., Peters, G. J., Abramson, R. D., Mullaly, S. J., Bosak, S. A., McEwan, P. J., McKernan, K. J., Malek, J. A., Gunaratne, P. H., Richards, S., Worley, K. C., Hale, S., Garcia, A. M., Gay, L. J., Hultk, S. W., Villalón, D. K., Muzny, D. M., Sodergren, E. J., Lu, X., Gibbs, R. A., Foley, J., Helton, E., Kettelman, M., Madan, A. C., Shvachko, Y., Sanchez, A. G., Whiting, M., Madan, A., Young, A. C., Shvachko, Y., Bonfield, G. G., Blakesley, R. W., Touchman, J. W., Green, E. D., Dickson, M. C., Rodriguez, A. C., Grimwood, J., Schwartz, J., Myers, R. M., Butler, M. D., Krzywicki, M. I., Skalska, U., Small, D. E., Schnerch, A., Schein, J. E., Jones, S. J., and Marra, M. A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (26)	16899-16903 (2002)
MEDLINE	23388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 2435)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-DEC-2002)		
	National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgabps-remail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC) Gatherersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nih.gov		
	Akhtar, N., Ayele, K., Beckstrom-Sternberg, S. M., Benjamin, B., Blakesley, R. W., Bonfield, G. G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N. L., Granito, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S. L., Karlins, E., Kwong, P., Latic, P., Legesky, R., Maduro, Q. L., Mastello, C., Maskeri, B., Mastrian, S. D., McCluskey, J. C., McDowell, C., Pearson, R., Steinert, S., Thomas, P. J., Touchman, J. W., Tsugeon, C., Vogt, J. L., Walker, M. A., Weethey, K. D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E. D.		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILMI.ac: <http://Image.liml.gov>
Series: IRAC Plate: 82 Row: d Column: 4.
Location/Qualifiers
1. .2435
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:47884 IMAGE:5194649"
/tissue_type="Brain, adult, 6 pooled whole brains"
/clone_id="NH_MGC_114"
/lab_host="DH10B"
/notes="Vector: pCMV-Sport6"
1. .2435
/gene="VMD2"
/notes="synonyms: BMD, BEST"
/db_xref="locusID:7439"
/db_xref="MIM:153700"
87. .1901
/gene="VMD2"
/codon_start=1
/product="VMD2 protein"
/protein_id="AAH41664.1"
/db_xref="GI:27371320"
/db_xref="locusID:7439"
/db_xref="MIM:153700"
/translation="MFKEKLTLYCDYSYIOLIPISFVLGFVYLVVTRWMNVEYNLPMPP
RLMSLVSGFVEGKDEQGLRLRLRLRYANLNLVLSVSVAVYKPPSAOHLVQAGF
MTPAEHKOLEKLSLPNNHFWPMWPMFLNLSKAWLGGRIIRPILLOSLNEMNTPTRO
CGHLAYADWISIPLYIVQVTVAVYSPFLCIVGRLQRLNPAPKAPRGHELDLVYVPTFF
LQFFVYWGMLVAQALINPREGDDDPETWNIYVRNLQVSLAYDENHODLPRMEPDM
YNNRPERPPYTPAASQPRFVSFEGSTFNISLNKEEMFQPNQDEBDAAGLIGRPL
GLQSDHHPFRANSRTKLMPKRSLHLEGLPKNHKAQVNRQEDNKAMKLAVADA
FKSAVLVQRPQGYAPQPIPLSPMPFPLPEPSAPKLSHYGIDTKDSEKTVSSGAF
KSPFLSLSDGALMEHPEVSSQRRKYTEFNITDMPRLPENHLKEPLKSPKNTHTTLLK
DHMPYMLAENRSLVHLNNOGCIACPTPALSIALPLPILNPLGFIHHCQSTLDRPAL
AMGYLVAFTGIILKCGSGFLTPSWNHPDPLGGBGR"
90. .1022
/gene="VMD2"
/notes="DUF289; Region: Putative membrane protein. This
family called family 8 in, may be a transmembrane protein
The specific function of this protein is unknown"
/db_xref="CD:pfam01062"

misc_feature
Query Match 56.4%; Score 1371; DB 9; Length 2435;
Beet Local Similarity 88.4%; Pred. No. 0;
Matches 1587; Conservative 0; Mismatches 5; Indels 203; Gaps 1;

ORIGIN

```

257 GCTGGCCCTCAGGAGAAACAACAGCTGAGTGTGAGAACTGACTGTGATTGGCAG 316
Db 59 GCTGGCCCTCAGGAGAAACAACAGCTGAGTGTGAGAACTGACTGTGATTGGCAG 118
317 CTACATCAGTCTATCCCACTTTCCTTGGTGTGGGCTTTACGTAGCGTGTGTGAC 376
Db 119 CTACATCAGTCTATCCCACTTTCCTTGGTGTGGGCTTTACGTAGCGTGTGTGAC 178
377 CCGCTGTGTGACACAGTACGAGAACCTGGCGGCGCGACCGGCTCATAGCGTGTGTG 436
Db 179 CCGCTGTGTGACACAGTACGAGAACCTGGCGGCGCGACCGGCTCATAGCGTGTGTG 238
437 GGGCTTGTGTAAGGACGAGACGAGCAAGGCGCGGCTGTGGCGCGACGCTCATCCGCT 496
Db 239 GGGCTTGTGTAAGGACGAGACGAGCAAGGCGCGGCTGTGGCGCGACGCTCATCCGCT 298
497 CGGCAACTTGGGGAAGGTGCTACTCTGTGGGACGGCTCAGACACCGCAGTCTCAAGGGCTT 556
Db 299 CGGCAACTTGGGGAAGGTGCTACTCTGTGGGACGGCTCAGACACCGCAGTCTCAAGGGCTT 358
557 CCCCAAGGCCACACACTGGTGCACAGAGCTTTATGATCTCCGGCGAAGACAAAGCATT 616
359 CCCCAAGGCCACACACTGGTGCACAGAGCTTTATGATCTCCGGCGAAGACAAAGCATT 418

```

Qy 617 GAGAACTGAGCCTACACAAATGTTCTGGGTGCTTGGTGTGTTGGCAACT 676
Db 419 GAGAACTGAGCCTACACAAATGTTCTGGGTGCTTGGTGTGTTGGCAACT 478
Qy 677 GTCAATGAGGCGTGGTGGAGTCCGAGCTTATCTGCTCCAGAGCTGCT 736
Db 479 GTCAATGAGGCGTGGTGGAGTCCGAGCTTATCTGCTCCAGAGCTGCT 538
Qy 737 GAACGAGTGAACACTTGGCTACTCACTGTGSAACCTGTATGCTTACGATTTAG 796
Db 539 GAACGAGTGAACACTTGGCTACTCACTGTGSAACCTGTATGCTTACGATTTAG 598
Qy 797 TATCCACTGTGTATACACAGTGTGACTGTGGCGTGTACAGCTTCTCTGACTTG 856
Db 599 TATCCACTGTGTATACACAGTGTGACTGTGGCGTGTACAGCTTCTCTGACTTG 658
Qy 857 TCTAGTTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGCGCATGAGCTGACCT 916
Db 659 TCTAGTTGGGCGGAGTTTCTGAACCCAGCCAGGCTTACCTGCGCATGAGCTGACCT 718
Qy 917 CGTTGTGCGGCTCTTCAAGTCTTCAAGTCTTCTATGTTGGCTGAGAGTGGG 976
Db 719 CGTTGTGCGGCTCTTCAAGTCTTCTATGTTGGCTGAGAGTGGGCTGAA----- 772
Qy 977 CCTCTCAGGCGCTGTGGCTGAGGCTAGGCGAGAGGCTATGGCCAGAGCTGCT 1036
Db 773 ----- 772
Qy 1037 TGAGACGAGATGCAAGTGTGAGAAAGAGTCTCAGCGGTAGAAAGACGACGCGTG 1096
Db 773 ----- 772
Qy 1097 GTGGCGCACACTGTATCCAGCTACTCGGAGAGCTGAGCAGAGAAATGCTTGAACC 1156
Db 773 ----- 772
Qy 1157 CGGAGCGGAGTGTGTGGAGAGAGCTCATCAACCTTTGGAGAGATGATGAT 1216
Db 773 -----GTTGGAGAGAGCTCATCAACCTTTGGAGAGATGATGAT 815
Qy 1217 GATTTGAGACCACTGATTTGTGACAGAAATTTGCAAGTGTCTCTGTTGGCTGTGAT 1276
Db 816 GATTTGAGACCACTGATTTGTGACAGAAATTTGCAAGTGTCTCTGTTGGCTGTGAT 875
Qy 1277 GAGATGACCAAGAGCTGCTCGATGAGAGCGGACATGTACTGGAATTAACCCGAGCA 1336
Db 876 GAGATGACCAAGAGCTGCTCGATGAGAGCGGACATGTACTGGAATTAACCCGAGCA 935
Qy 1337 CAGCCCCCTTACAGAGCTGCTTCCGCCAGTTCGTCGAGGCTCTTTATGGGCTCCACC 1396
Db 936 CAGCCCCCTTACAGAGCTGCTTCCGCCAGTTCGTCGAGTCTCTTTATGGGCTCCACC 995
Qy 1397 TTCAACATCAGCCTGAACCAAGAGAGATGAGTTCCAGCCCAATCAGAGAGAGAGAG 1456
Db 996 TTCAACATCAGCCTGAACCAAGAGAGATGAGTTCCAGCCCAATCAGAGAGAGAGAG 1055
Qy 1457 GATCTCAAGCTGGCATATTGGCCGCTTCTTAGGCTGCAAGTCCCATATCACAATCT 1516
Db 1056 GATCTCAAGCTGGCATATTGGCCGCTTCTTAGGCTGCAAGTCCCATATCACAATCT 1115
Qy 1517 CCCAGGCAAACTAAGAGCAAACTACTGTGGCCCAAGAGAGATCCCTTCTCAGAG 1576
Db 1116 CCCAGGCAAACTAAGAGCAAACTACTGTGGCCCAAGAGAGATCCCTTCTCAGAG 1175
Qy 1577 GAGCTGCCAAACCAACAGAGAGCAAGAGATGAGGAGCCAGAGAGCAACAG 1636
Db 1176 GAGCTGCCAAACCAACAGAGAGCAAGAGATGAGGAGCCAGAGAGCAACAG 1235
Qy 1637 GCTCTGAGAGCTTAAGGCTGTGAGAGCTTCAAGTGTGAGCCCACTGTATCAAGGCGAGGC 1696
Db 1236 GCTCTGAGAGCTTAAGGCTGTGAGAGCTTCAAGTGTGAGCCCACTGTATCAAGGCGAGGC 1295
Qy 1697 TACTACAGTGGCCACAGAGAGCCCTCAGCCCACTCCATGTTCTTCCCTTGAACA 1756

Db 1296 TACTACAGTGGCCACAGAGAGCCCTCAGGCCCACTCCCATGTTCTTCCCTTGAACA 1355
Qy 1757 TCAGCGCCCTTAAAGCTTCAACAGTGTCAACAGGCAATAGACACCAAGAAAGCTTAAG 1816
Db 1356 TCAGCGCCCTTAAAGCTTCAACAGTGTCAACAGGCAATAGACACCAAGAAAGCTTAAG 1415
Qy 1817 ACTGTAGTTCTGTGGGCGCAAGAAAGTTTGAATTGCTCTCAGAGAGGATGGGCTTGG 1876
Db 1416 ACTGTAGTTCTGTGGGCGCAAGAAAGTTTGAATTGCTCTCAGAGAGGATGGGCTTGG 1475
Qy 1877 ATGAGACACCCAGAGATATCTCAAGTAGAGAGAGAAACCTGTGAGTTTAACCTGAGAT 1936
Db 1476 ATGAGACACCCAGAGATATCTCAAGTAGAGAGAGAAACCTGTGAGTTTAACCTGAGAT 1535
Qy 1937 ATGCGAAGATCCCGGAAATACCTCAAGAAACCTTTGGAAACATACCAACCAACTA 1996
Db 1536 ATGCGAAGATCCCGGAAATACCTCAAGAAACCTTTGGAAACATACCAACCAACTA 1595
Qy 1997 CACACTACACTCAAGATCAATGATCTTATTTGGGCTTGGAAACAGGATG 2051
Db 1596 CACACTACACTCAAGATCAATGATCTTATTTGGGCTTGGAAACAGGATG 1650

RESULT 12
BD132659 1263 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD132659
VERSION BD132659.1 GI:23227604
KEYWORDS JP 2002504822-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1263)
AUTHORS Jacobs,K., McCoy,J.M., Lavallie,B.R., Racie,L.A., Treacy,M., Spaulding,V., Agostino,M.J., Howes,S.H. and Reachel,K.
TITLE Secreted proteins and polynucleotides encoding them
JOURNAL Patient: JP 2002504822-A 2 12-FEB-2002;
GENETICS INSTITUTE INC
PN JP 2002504822-A/2
PD 12-FEB-2002
PF 08-JUN-1998 JP 1999503038
PR 11-JUN-1997 US 08/873218, 05-JUN-1998 US 09/092722 PI
KENNETH JACOBES,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
MAURICE TREACY,
PI VIKKI SPAUDLING,MICHAEL J AGOSTINO,STEVEN H HOWES,KIM PECHTEL,
PC C12N15/12,C07K14/47,A61K38/17
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source location/Qualifiers
1..1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 51.2%; Score 1244.2; DB 6; Length 1263;
Best Local Similarity 99.8%; Pred. No. 2.2e-284;
Matches 1246; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1181 GAGCAGTTCATCAACCCCTTTGAGAGAGATGATGATTTTGGAGCCAACTGATTTGTC 1240
Db 1 GAGCAGTTCATCAACCCCTTTGAGAGAGATGATGATTTTGGAGCCAACTGATTTGTC 60
Qy 1241 GACGAGATTTGCAAGTGTCTCTGTTGGCTGTGATGAGATGACACGAGACCTGCTCGG 1300
Db 61 GACGAGATTTGCAAGTGTCTCTGTTGGCTGTGATGAGATGACACGAGACCTGCTCGG 120
Qy 1301 ATGAGCGGAGCATGTACTGGAATTAAGCCGAGCAGAGCCCTTACACAGCTGCTTCC 1360


```

Db      121 ATGAGCCGGACATGATCTAGTAATAGCCCGAGCCAGACCCCTTACAGACTGCTTCC 180
Oy      1361 GCCCAGTTCCGTCGAGCCCTCTTATGAGCTCCACTTCAACATAGAGCTGAACAAAG 1420
Db      181 GCCCAGTTCCGTCGAGCTCTTATGAGCTCCACTTCAACATAGAGCTGAACAAAG 240
Oy      1421 GAGATGAGTTCCAGCCCAATCAGAGAGAGAGAGAGATGCTCAGCTGAGCATCTTGGC 1480
Db      241 GAGATGAGTTCCAGCCCAATCAGAGAGAGAGAGATGCTCAGCTGAGCATCTTGGC 300
Oy      1481 GCCTTCTTAGGCTCTGACATCCCATATATCACTCTCCAGGGCAAATCTCAAGACCAA 1540
Db      301 GCCTTCTTAGGCTCTGACATCCCATATATCACTCTCCAGGGCAAATCTCAAGACCAA 360
Oy      1541 CTACTGTGAGCCCAAGGGAATCCCTTCTCAAGAGGGCTGCCCCCAACCAAGGCA 1600
Db      361 CTACTGTGAGCCCAAGGGAATCCCTTCTCAAGAGGGCTGCCCCCAACCAAGGCA 420
Oy      1601 GCCAAACAGAACTTTAGGGGCGCAGAGAGACAAAGGCTGGAAGCTTAAAGCTGTGAC 1660
Db      421 GCCAAACAGAACTTTAGGGGCGCAGAGAGACAAAGGCTGGAAGCTTAAAGCTGTGAC 480
Oy      1661 GCCTTCAAGTCTGCGCCCACTGTATAGAGGCGCAGCTCTACAGTGCCTCAGAGCC 1720
Db      481 GCCTTCAAGTCTGCGCCCACTGTATAGAGGCGCAGCTCTACAGTGCCTCAGAGCC 540
Oy      1721 CTCAGCCCCACATCCCATGTTCTTCCCTTAGAACCATCAGGCGCTCAAGCTTCAAGT 1780
Db      541 CTCAGCCCCACATCCCATGTTCTTCCCTTAGAACCATCAGGCGCTCAAGCTTCAAGT 600
Oy      1781 GTCAAGGCAATAGACCAACAAAGAAAGCTTAAAGACTGTGAGTCTGGGGCCAGAAA 1840
Db      601 GTCAAGGCAATAGACCAACAAAGAAAGCTTAAAGACTGTGAGTCTGGGGCCAGAAA 660
Oy      1841 AGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1900
Db      661 AGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Oy      1901 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1960
Db      721 GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Oy      1961 CTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2020
Db      781 CTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Oy      2021 GATCTCTTATGAGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2080
Db      841 GATCTCTTATGAGGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Oy      2081 GGATGCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2140
Db      901 GGATGCTTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Oy      2141 CAGGCAATACAGCTGTCCACTGAAAGAAAGCTGTCTTCAACAGGCTGAATCAATGCTTA 2200
Db      961 CAGGCAATACAGCTGTCCACTGAAAGAAAGCTGTCTTCAACAGGCTGAATCAATGCTTA 2260
Oy      2201 GCTTAAATAGATTAATAATCCAGACTACTTCAAGCCTTTAAAGCTTTTATTAATAAAT 2260
Db      1021 GCTTAAATAGATTAATAATCCAGACTACTTCAAGCCTTTAAAGCTTTTATTAATAAAT 1080
Oy      2261 GTGAAAGCTAGACTGAACCAATGGAACATTTAATCAGACTGTGATTCAGAGTCCGGA 2320
Db      1081 GTGAAAGCTAGACTGAACCAATGGAACATTTAATCAGACTGTGATTCAGAGTCCGGA 1140
Oy      2321 ACCCTTAGTCTATCTGAATCAAGACCAACCTTAGTATCTGCCAAACTAATGA 2380
Db      1141 ACCCTTAGTCTATCTGAATCAAGACCAACCTTAGTATCTGCCAAACTAATGA 1200
Oy      2381 GTTAAATTAATTAATCAATCTCGTTAAATTAATTAATTAATTAATTAATTAATTAAT 2429
Db      1201 GTTAAATTAATTAATCAATCTCGTTAAATTAATTAATTAATTAATTAATTAATTAAT 1249

```

```

RESULT 13
BC079048
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 2035)
Strasbourg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Bueltel,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heile,F.,
Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Adamsom,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalske,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2035)
Director MGC Project.
Direct Submission
Submitted (02-AUG-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cga@ncl.nih.gov
Tissue Procurement: Howard Jacobs
CDNA Library Preparation: Express Genomics
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.bhg.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRK Plate: 183 Row: a Column: 22
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
Location/Qualifiers
1..2035
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:93985 IMAGE:7115295"
/rissue_type="Testis, rat (Brown Norway)"
/clone_id="NIH_MGC_237"
/lab_host="DH10B"
/notes="Vector: pExpress1"

```

FEATURES

source

CDs	153. .1805	/codon_start=1
	/product="Unknown (protein for MGc.93985)"	
	/protein_id="AAH79048.1"	
	/db_xref="GI:50927693"	
	/translation="MTITTYNEVANARLIFSFCLLLRMSGIYKLYGEPLVIFPLY	
	CIKRLYRMVLVSNHQOQPEKLALYCNSTYQIPISFVLGFTVLIVSMNQYENLMP	
	PDRLMVOVSFVEGEDEQRLRLTILYALIGQVLIIRSISTVYSRPSLOHLMOP	
	GFMTPEKXQLORLGLPNTFPMVPMVMVAMSMKAYLDGRIIDTVLYLOSLENEVCTLR	
	TOGHLVAYDKMINIPLYTVYTVYVAFIACILGRFPLAPLKNRBEHMDLYVPVE	
	TILOPLFTWGLKVABQILNPFGEDDDFETNWIIDRLQVLSVDSMHDVPLEP	
	DLVYNEAAPQPPYTAASGSHRSQSFMSFTMSLKKDSEESDITYDKENDYSSITG	
	RFLGLOSNDYHPKTDLKTKLMSKQPFLEDSKNDQKQKNDVMPKDLFPAIIPK	
	FKRPSYRQAPQPLSHLPMPVPPPOOASVSYDPSGSAQYQAEACMKKKYVEFNL	
	NIPRCPIHEBPHLDQHTILKEPESIFPVKADTNVSVYK"	
ORIGIN		
Query Match	29.3%; Score 712.8; DB 10; Length 2035;	
Beet Local	Similarity 67.9%; Pred. No. 3.8e-158;	
Matches 1182;	Conservative 0; Mismatches 327; Indels 233; Gaps 4;	
Qy	50 CCCACCTGGAACCCACCTGACCCCAAGCCCACTGCTGACGCCCATGCTGGCCATGAC	109
Db	98 CCAACAGAAACCCCTAACAGAGACCCCAAGCCCAAGCTCTACTGATGCAAGGCATGAC	157
Qy	110 CATCACTTACACAAGCCCAAGTGGCTAATGCCGCTTAAAGCTCTTCGCCGCTGCT	169
Db	158 TATCACTTACACAAGCCCAAGTGGCCCAATGCCGCTTAAAGCTCTTCGCCGCTGCT	217
Qy	170 GTGCTGGCGGGAGCATCTACAAAGCTCTAATGAGCAGTTCCTTATCTTCCTGCTG	229
Db	218 GCGCTGGGAGGAGCATCTACAAAGCTCTGATGAGGAATTCCTTATCTTCATATTCCT	277
Qy	230 CTACTACATCATCCGCTTATTTATAGGCTGGCCCTTCACGGAAGAACAGCTGATTT	289
Db	278 CTACTATTCATCCCTGACTCTACAGATGTTCTCTCGAATCATGACAGACAGCATTT	337
Qy	290 TGAGAACTGACTCTGATTTGGAGCAGCTATCCAGCTATCCCATTCCTTCCTGCT	349
Db	338 TGAGAACTGACTCTGATTTGGAGCAGCTATCCAGCTATCCCATTCCTTCCTGCT	397
Qy	350 GGGCTTCTACGTGACGCTGCTGTGACCCGCTGTGTGAAACAGTACGAAACCTGCCGTG	409
Db	398 GGGTTTCTATGTTACATTTGTGTGTGAGGCGAATGTTGGAACAGTACGAAACCTGCCGTG	457
Qy	410 GCCCGACCGCTCATGAGCCTGTGTGCGGGCTTGTGCAAGGCAAGACGAGCAAGCGCG	469
Db	458 GCCCGACCGCTCATGAGCCTGTGTGCGGGCTTGTGCAAGGCAAGACGAGCAAGCGCG	517
Qy	470 GCTGCTGGGGGCAAGCTCATCCGCTAGGCCCACTGGGGCAAGCTGCTCATTCCTGGGAG	529
Db	518 GTTGTGGGGGCAAGCTCATCCGCTAGGCCCACTGGGGCAAGCTGCTCATTCCTGGGAG	577
Qy	530 CGTCAGCACCCGAGCTCTACAGCGCTTCCCGACGCCAGCACTGTGTGCAAGCAGGCTT	589
Db	578 CATCAGCACCTTCGCTACAGCGCTTCCCGAGCTTTCACACCTGATGCAAGCAGGCTT	637
Qy	590 TATGATCTCGGCAAGAACAGAGCTGAGAAACTGAGCTTACCAACAACATGTTCTG	649
Db	638 TATGATCTCGGAGAAATTAAGCAGTTGACAGGGTTGGGCTTACCAACAACAATTTCTG	697
Qy	650 GGTGCGCCGGGTGTGTGTTGCAACCTGTCAATGAAAGCGGTGTGGAAGGTCCAAATCCG	709
Db	698 GGTACCCCTGGGTGTGTGTTGCAACCTGTCAATGAAAGCTTATCTTGAAGGTGCATATAG	757
Qy	710 GGAACCTTATCTGCTCCAGAGCCTGTGAAAGAGATGAAACCTTGTCTACTAGTGTG	769
Db	758 GGAACCTTATCTGCTCCAGAGCCTGTGAAAGAGATGAAAGCTTGTCTACTAGTGTG	817
Qy	770 AACCTGTATGCTTACGACTGATTAATGATTCCTGCTGTATATACAGAGTGTGATCTGT	829
Db	818 AACCTGTATGCTTATGACTGATTAATGATTCCTGCTGTATATACAGAGTGTGATCAGT	877

Qy	830 GGCAGGTGACAGCTTCTTCTGACTTGTCTAGTGGGCGGAGTTTCTGAACCCAGCCAA	889
Db	878 GGCAGGTTACAGCTTCTTCTTCTGATGCTGATCGGAGGAGGATTTCTGAACCCAGCTCA	937
Qy	890 GGCCTTACCTTGGCCATGAGCTGAGACCTGTGTGTGCGCGCTTTACAGTTCCTGCAATTC	949
Db	938 GAACTACCCAGGCCACGAGATGAGATCTGTGTGTGCTGTCTTCAAACTCTGCAATTCCT	997
Qy	950 CTTCTATGTTGGCTGGCTGGAAGGTGGGCTCTCCAGGGCCCTGCTGGGCTGAGGATATG	1009
Db	998 ATTCTACATGGCTGGCTGCA-----	1018
Qy	1010 CCAGAGGGGTGATGAGCCAGAGCTGTGAGACGAGATGACAGTGTCAAGAAAGAGAGT	1069
Db	1019 -----	1018
Qy	1070 CTCACGGGTAGAAAGACAGCCAGCGTGTGTGGCGCACACTGTAAATCCAGCTACTCGGA	1129
Db	1019 -----	1018
Qy	1130 GGCAGAGCAGAGAGATTCCTTGAACCCGGGAGGGAGGTTGTGTGGCAGAGAGCTC	1189
Db	1019 -----	1018
Qy	1190 ATCAACCCCTTGGAGAGAGATGATGATATTTGAGACCAACTGATTTGTGACAGAGAT	1249
Db	1035 ATCAACCCCTTGGAGAGAGATGATGATATTTGAGACCAACTGATTTGTGACAGAGAT	1094
Qy	1250 TTGCAAGTGTCCCTGT	1309
Db	1095 CTGCAAGTGTCCCTGT	1154
Qy	1310 GACATGTACTGGAATTAAGCCCGGAGCCAGCCGCCCCCTTACAGAGCTTCCGCCAGTTC	1369
Db	1155 GACCTGTACTGGAATTAAGCCCGGAGCCGCCCCCTTACAGAGCTTCCGCCAGTTC	1214
Qy	1370 CGTGAAGCTCTTATATGAGGCTCCACCTTCAACATCAGCTGAAACAAAGAGAGATGAG	1429
Db	1215 CACGAGCAATCTTCTATGT	1274
Qy	1430 TTCAG---CCCAATCAGAGAGAGAGAGAGATCTCAGCTGCGATCTTGGCCGCTTC	1486
Db	1275 GAGGAGATGTACACTTATGACAAAGAAAGATATTAACGAGACACCATTTGAGACGCTTC	1334
Qy	1487 CTAGGCTGTGAGTGTCCATGATCAACATCTCTCCAGGGGAACTCAAGAACCAAACTACTG	1546
Db	1335 TTAGAGACTGCAATCAATGATCAATCTCTCCAGAGAGACTTAAAGAACCAAACTACTG	1394
Qy	1547 TGCGCCAAAGAGAGATCCCTTCTCCAGAGGGCTGTGCCAAAGAACCAAGAGCAGCCAA	1606
Db	1395 TGCTGTAGAGAG---CCCTTCTGTGAGGACCAAGTCTTAAGAAATGACAG-----	1442
Qy	1607 CAGAACTTAAAGGAGCAGAGAAAGCAAGAGCTTGAAGCTTAAAGCTGTGTGAGCGCTTC	1666
Db	1443 -----CAGAAAGACAAAGAAATATGATGTCTGAAACCAAGAGACTCTTC	1487
Qy	1667 AAGTCTGCCCACTGTATCAGAGGCGAGGCTTACTACAGTGTGCCCCACAGACGCCCTCAGC	1726
Db	1488 AAGCTTATACCAAGATTTAAAGAGGCCAAATGACCGCCAGGCCCCACAGCCACTTGTGAGC	1547
Qy	1727 CCACTCCCATGTTCTTCCCTTGAAGCAATCAGAGCGCGGTCAAAAGCTTCAAGATGTACA	1786
Db	1548 CACCTTCCCATGTGTCTTCCCTTCAAGCAGTCAAGACCCCTCAGATGTCTTACTCAGACACA	1607
Qy	1787 GG 1788	
Db	1608 GG 1609	
RESULT 14		
AY450427 1904 bp mRNA linear ROD 24-JUN-2004		
LOCUS AY450427		
DEFINITION Mus musculus vitelliform macular dystrophy 2 (Vmd2) mRNA, complete		

Db 1118 GCTCCACCTTCAACATCAGCCTTAAGAAAGAAAGACTTGAAGCTTGTGTCAAAAGAGAGG 1177
Qy 1449 ACGAAGAGATGCTCACGCTG-----GCATCATTTGGCGGCTTCTAGGCGTCGC 1496
Db 1178 CTGACACGGATTAAGAAAGAGTGGCTATAGCAGACCATATAGGCTGCTTTTAAAGACATGC 1237
Qy 1497 AGTCCCATGATCACCATCTCTCCAGGGCAAACTCAAGAGCAAACTACTGTGGCCCAAGA 1556
Db 1238 AACCCAAAATACATCTCTCTGAAAGACTTAAAGACCAAACTATTTGTTCTAAGA 1297
Qy 1557 GGGAAATCCCTTCTCCAGAGGGCTGCCCCAAAACCAAGAGCAGCCAAACGAACGTTA 1616
Db 1298 -----ACCCCTCTCCAGAGCCAGTGTAAAGATGCCAACGAAAAACGAAG--- 1348
Qy 1617 GGGGACAGAGAGAACAAAGGCTGAGACTTAAGAGCTGTGAGAGCCTTCAAGTCTGGCC 1676
Db 1349 -----ATGCTGGAAATTTAAAGGCTCTGAGCTTCTTGAATGTGTTC 1390
Qy 1677 CACTGTATCAGAGGCGCAGCTACTACAGTGCCTCCACAGACGCGCCCTCAGCCCACTCC 1734
Db 1391 CAAGTTTAAAGAGAGAGGCTCCCATTTGTGGCCCAAGAGCAACCCAGAGCCACCTTAC 1448

RESULT 15
BD136744
LOCUS Beet's macular dystrophy gene. 1916 bp DNA linear PAT 18-SEP-2002
ACCESSION BD136744
VERSION BD136744.1 GI:23231689
KEYWORDS JP 2002504559-A/26.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1916)
Petrushkin, K., Caskey, T. C., Metzker, M. and Wadelius, C.
Beet's macular dystrophy gene
Patent: JP 2002504559-A 26 12-FEB-2002;
JOURNAL MERCK & CO INC, CLAES WADELIVS
COMMENT OS Mus musculus (mouse)
PN JP 2002504559-A/26
PD 12-FEB-2002
PF 22-FEB-1999 JP 2000533447
PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUSHKIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI
WADELIVS
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC
C12P21/08,
PC C12N5/00, C12N15/00
CC Beet's macular dystrophy gene
FH key Location/Qualifiers
FT source 1..1916
FT Location/Qualifiers
1..1916
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match 27.1%; Score 658.2; DB 6; Length 1916;
Beet Local Similarity 67.6%; Pred. No. 3.5e-145; Indels 242; Gaps 4;
Matches 1116; Conservative 0; Mismatches 293; Indels 242; Gaps 4;

Db 1122 GTCTTCATATTCCTCTCTATTCATCTCGTAGCTTACAAATGTTCTCTCAGTGTAT 181
Qy 276 CAACAGCGATGTTTGAAGAACTGACTGTGATTTGGCAGACTCATCAGCTCATCC 335
Db 182 CAGACGCTGTGTTGAAGAGCTGCTGTGATCGCAGACTATCATCAGTCTATCCCT 241
Qy 336 ATTTCTTCTGTGTGGGCTTCTACGTGACGTGTGTGTGACCCGCTGTGTGAACCAATAC 395
Db 242 ATATCTTCTGTGTGGGCTTCTATGTATCATTTGATGTGTGACCCGCTGTGTGAACCAATAC 301
Qy 396 GAGAACTGCGGTGGCCCGACCGCTCATAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 455
Db 302 GAGAACTTGTGCGGTGGCCCGACCGCTCATGTATCCAGGTGTGTGTGTGTGTGTGTGTGTGT 361
Qy 456 GAGCAGAAAGGCGGCTGT 515
Db 362 GATGAGAAAGGCGGCTGT 421
Qy 516 CTCACTCTGCGCAGCTGTGACAGCGAGTCTACAAAGCGCTTCCCGAGCGCCAGCAGCTG 575
Db 422 CTCACTCTGCGCAGCTGTGACAGCGAGTCTACAAAGCGCTTCCCGAGCTTCCAGCAGCTG 481
Qy 576 GTGCAAGAGGCTTTATGACTCCGCAAGAACAAAGAGTTGAGAACTGAGCTTACCA 635
Db 482 GTGCTAGAGAGTTTATGACCATGGGGAACATAGCAGTTGAGAAAGTTGGGCTTACCA 541
Qy 636 CACAACATGTTCTGGGTGCGCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 695
Db 542 CACAACATGTTCTGGGTGCGCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 601
Qy 696 GGAGGTGAATCCGGGAGCCCTATCTGTCTCCAGAGCCGTGAGCGAGATGAACACTTGG 755
Db 602 GGAGGTGAATCCGGGAGCCCTATCTGTCTCCAGAGCCGTGAGATGAATGAGGTGTGACTTGG 661
Qy 756 CGTACTAGTGTGACACCTGTATAGCTTACAGCTGATTAAGTATCCCATGCTGTATACA 815
Db 662 CGTACTAGTGTGACACCTGTATAGCTTACAGCTGATTAAGTATCCCATGCTGTATACA 721
Qy 816 CAGGTGTGACTGTGGGCTGTACAGCTTCTTCTGACTGTGTGTGTGTGTGTGTGTGTGTGTGT 875
Db 722 CAGGTGTGACTGTGGGCTGTACAGCTTCTTCTGACTGTGTGTGTGTGTGTGTGTGTGTGTGT 781
Qy 876 CTGAACCAAGCAGAGGCTTCCCGCAGATGAGTGTGACCTGTGTGCGGCTGTACAG 935
Db 782 CTGAACCAAGCAGAGCTTCCCGCAGATGAGTGTGACCTGTGTGCGGCTGTGTACAG 841
Qy 936 TTCTGTGAGTTCTTCTTATGT 995
Db 842 ATCTGTGATTTCTTATTTCTATATGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 876
Qy 996 GCGTGAAGCATGGCCAGAGGGTTCATGCGCAGACGCTGTTGAGACGAGGATGCAGTGT 1055
Db 877 ----- 876
Qy 1056 CAGGAAGGAAGTGTCTACGCGGTGAAGAACGACCGCGTGTGGCCGACACTGTATC 1115
Db 877 ----- 876
Qy 1116 CCACTACTCGGAGGCTGAGGAGCAGAGAAATCGCTTGAACCCGGAGGCGAGGTTGTGG 1175
Db 877 -----GG 878
Qy 1176 TGGCAGAGCAGCTGATCAACCCCTTGGAGAGATGATGATGATTTGAGCAACTGGA 1235
Db 879 TGGCAGAGCAGCTGATCAACCCCTTGGAGAGAGATGATGATGATTTGAGCAACTGGA 938
Qy 1236 TTGTGACAGGAATTTGCAAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1295
Db 939 TCATTTGACAGAAACCTGCAAGGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 998
Qy 1296 CTGCGATGAGGCGGAGATGTACTGAAATAAGCCGAGGCCAGACCCCTTACAGAGCTG 1355

```
Db      999  CTCCCATGGAACGTGACATGTACTGGAACGAGGCAAGCGCTCAGCCGCCCTACACAGCTG 1058
QY      1356  CTTCCGCCCAAGTTCGCTGAGACCTCCTTATGAGGCTCCACTTCAACATGAGCCTGAACA 1415
Db      1059  CTTCTGCGAGGTCTCGCGGCACTTCCTTCATGAGGCTCCACTTCAACATGAGCCTGAAGA 1118
QY      1416  AAGAGGATGAGGTTCAGGCCCAATCAGAGGACGAGAGATGCTCACGCTG----- 1469
Db      1119  AAGAAGACTTAGACTTTGTCAAAAGAGAGGCTGACAGGATTAAGAAAGAGGTGGCT 1178
QY      1470  -----GCATCATTTGGCCGCTTCTCTAGGCTTGCAGTCCATGATCACCATCTCCAGGG 1523
Db      1179  ATAGCAGACCATAGGCTGCTTCTTAGGACTGCAACCCAAAACCTACCATCTTCCCTTGA 1238
QY      1524  CAAACTGAAGGACCAACTACTGTGGGCCCAAGAGGAAATCCCTTCTCCAGAGGGGCTGC 1583
Db      1239  AAGACTTAAAGACCAAACTATTGTGTTTAAAGA-----ACCCCTCTCTCGAAGGCCAGT 1292
QY      1584  CCAAAAAACACAGAGCAGCCAAACAGAACGTTAGGGGCCAGAAAGACAAACAGGCTTGA 1643
Db      1293  GTAAGATGCCAACCAAGAAAACAGAAAG-----ATGCTTGA 1331
QY      1644  AGCTTAAAGCTGTGAGCGCCTTCAAGTGTGGCCCACTGTATCAGAGGCCAGGCTACTACA 1703
Db      1332  AATTTAAGGGGTCTGACTTCTTGAAATGTGTCCAAAGGTTTAAAGAGAGGCTCCCATTT 1391
QY      1704  GTGCCCCACAGAGCGCCCTCAGCCCACTCC 1734
Db      1392  GTGGCCCAAGGCAACCCAGAGCCACCTTAC 1422
```

Search completed: March 26, 2005, 07:38:57
Job time : 10382.6 secs

double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M. Patricia Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."

ORIGIN

Query Match 28.1%; Score 681.8; DB 5; Length 711;
Best Local Similarity 99.6%; Pred. No. 1.9e-137;
Matches 694; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1536 CCAACTACTGTGTGCCCCAG-AGGGAATCCCTTCTCCAGAGGCGCTGCCAAAAACAC 1594
DB 1 CCAAACTACTGTGTGCCCCAGAGNAGGAATCCCTTCTCCAGAGGCGCTGCCAAAAACAC 60
QY 1595 AAGGACGCCAAAGAAAGCTTAAAGGGCCAGAGAAACAAGGCGCTGGAAGCTTAAAGCT 1654
DB 61 AAGGACGCCAAAGAAAGCTTAAAGGGCCAGAGAAACAAGGCGCTGGAAGCTTAAAGCT 120
QY 1655 GTGAGCGCTTCAAGTCTGAGCCCACTGTATCAGAGGCGAGGCTACTAGAGTCCACAG 1714
DB 121 GTGAGCGCTTCAAGTCTGAGCCCACTGTATCAGAGGCGAGGCTACTAGAGTCCACAG 180
QY 1715 AGCGCCCTCAGCGCCCACTCCCATGTCTTCTCCCTTGAACCATCAGCGCGCTCAAGCTT 1774
DB 181 AGCGCCCTCAGCGCCCACTCCCATGTCTTCTCCCTTGAACCATCAGCGCGCTCAAGCTT 240
QY 1775 CACAGTGTACAGAGGCAATACACCAAGAAAGCTTAAAGCTGATGATCTTGGGCGC 1834
DB 241 CACAGTGTACAGAGGCAATACACCAAGAAAGCTTAAAGCTGATGATCTTGGGCGC 300
QY 1835 AAGAAAGTTTGAATGTCTTCAAGAGCGATGGGCGCTTGATGAGAGCCCAAGAGTA 1894
DB 301 AAGAAAGTTTGAATGTCTTCAAGAGCGATGGGCGCTTGATGAGAGCCCAAGAGTA 360
QY 1895 TCTCAAGTGAAGAGAAAGCTGAGAGTTAACTGACGAGATATCCAGAGATCCCGCA 1954
DB 361 TCTCAAGTGAAGAGAAAGCTGAGAGTTAACTGACGAGATATCCAGAGATCCCGCA 420
QY 1955 AATCACTTCAAGAAAGCTTTGGAACATACCAACCAATACATACATACATCACTAAGAT 2014
DB 421 AATCACTTCAAGAAAGCTTTGGAACATACCAACCAATACATACATACATCACTAAGAT 480
QY 2015 CACATGATCTTATTTGGGCGCTTGAAGAAAGAGGATGAAGACATCTCTAAGCTTCC 2074
DB 481 CACATGATCTTATTTGGGCGCTTGAAGAAAGAGGATGAAGACATCTCTAAGCTTCC 540
QY 2075 TAATGGAGATGCTTGGCAGAGGATCTCTCACTGTGTATACACAGAGAGACATGATC 2134
DB 541 TAATGGAGATGCTTGGCAGAGGATCTCTCACTGTGTATACACAGAGAGACATGATC 600
QY 2135 CAGTCAAGCCATACAGTGTTCACACTGAAGAAAGTGTCTTACAGAGCGCTGAATCAA 2194
DB 601 CAGTCAAGCCATACAGTGTTCACACTGAAGAAAGTGTCTTACAGAGCGCTGAATCAA 660
QY 2195 TGGTTAGTTAATGATTAATAATCCCACTACTTCA 2231
DB 661 TGGTTAGTTAATGATTAATAATCCCACTACTTCA 697

RESULT 2

BU741926 676 bp mRNA linear EST 10-OCT-2002
LOCUS BU741926
DEFINITION UT-E-B01-aj-a-j-23-0-UI.s1 UT-E-B01 Homo sapiens cDNA clone
ACCESSION BU741926
VERSION BU741926.1 GI:23687736
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 676)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
JOURNAL discovery
MEDLINE Genome Res. 6 (9), 791-806 (1996)
PUBMED 97044477
8889548

COMMENT

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9585
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

source

Location/Qualifiers
1..676
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-B01-aj-a-j-23-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-B01"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site: 1: EcoR I; Site: 2: Not I;
UT-E-B01 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CGCGTATACC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI).
TAG_TISSUE=human fetal eye
TAG_LIB=UI-E-B01
TAG_SEQ=CGCGTATACC"

ORIGIN

Query Match 27.0%; Score 657; DB 5; Length 676;
Best Local Similarity 99.1%; Pred. No. 4.6e-132;
Matches 671; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1746 CCTTGAACCATCAGCGCGCTCAAGCTTCAAGTGTACAGGATAGACCCAAAGACA 1805
DB 676 CCTTGAACCATCAGCGCGCTCAAGCTTCAAGTGTACAGGATAGACCCAAAGACA 618
QY 1806 AAAGCTTAAAGACTGTGATTTCTGGGCGCAAGAAAGTTTGAATTCCTCAGAGAGCG 1865
DB 617 AAAGCTTAAAGACTGTGATTTCTGGGCGCAAGAAAGTTTGAATTCCTCAGAGAGCG 558
QY 1866 ATGGGCGCTTATGAGAGACCCCAAGATATCTCAAGTGAAGAGAAACTGTGAGTTTA 1925
DB 557 ATGGGCGCTTATGAGAGACCCCAAGATATCTCAAGTGAAGAGAAACTGTGAGTTTA 498
QY 1926 ACCTGACGATATGACAGAGATCCCGAAATACCTCAAGAAACCTTGGAAACATCAC 1985
DB 497 ACCTGACGATATGACAGAGATCCCGAAATACCTCAAGAAACCTTGGAAACATCAC 438

QY 1986 CAACCAATACACACTACTCAAAAGATCATAGATCTTATGGCTTGGAAAAA 2045
DB 437 CAACCAATACACACTACTCAAAAGATCATAGATCTTATGGCTTGGAAAAA 378
QY 2046 GGGATGAAGACATCTCTACCTGCTTCTTAATGGGGATGCTTGGCCAGCAAGTCTCA 2105
DB 377 GGGATGAAGACATCTCTACCTGCTTCTTAATGGGGATGCTTGGCCAGCAAGTCTCA 318
QY 2106 CCTGTGTATACACACAGACAGATCATAGATCTTCAACAGCAAGCTTCCACACTGAA 2165
DB 317 CCTGTGTATACACACAGACAGATCATAGATCTTCAACAGCAAGCTTCCACACTGAA 258
QY 2166 GAACGTGTCTTCAACACAGCTTGAATCAATGATGTTAATGATTAATAATCCAGACT 2225
DB 257 GAACGTGTCTTCAACACAGCTTGAATCAATGATGTTAATGATTAATAATCCAGACT 198
QY 2226 ACTTCAGCTTTAATGCTTTTATTCATTAATAAAGTGAAGCTGAAAGCTGAAAGCTTGA 2285
DB 197 ACTTCAGCTTTAATGCTTTTATTCATTAATAAAGTGAAGCTGAAAGCTGAAAGCTTGA 138
QY 2286 AACATTTAATCAACACTGATCTGATTCAGAGTGGGAAACCTTATGTTCTATCTGAATCCAG 2345
DB 137 AACATTTAATCAACACTGATCTGATTCAGAGTGGGAAACCTTATGTTCTATCTGAATCCAG 78
QY 2346 AACACCAACCTTATGATCTGATTCAGAGTGGGAAACCTTATGTTCTATCTGAATCCAG 2405
DB 77 AACACCAACCTTATGATCTGATTCAGAGTGGGAAACCTTATGTTCTATCTGAATCCAG 18
QY 2406 AAAAAAAAAAAAAAAAAA 2422
DB 17 AAAAAAAAAAAAAAAAAA 1

RESULT 3
BU730894/c 659 bp mRNA linear EST 09-OCT-2002
LOCUS UI-E-C11-afw-1-11-0-UI.s1 UI-E-C11 Homo sapiens cDNA clone
DEFINITION UI-E-C11-afw-1-11-0-UI 3', mRNA sequence.
ACCESSION BU730894
VERSION BU730894.1 GI:23655244
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 659)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLY-A=yes
Location/Qualifiers
1..659
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C11-afw-1-11-0-UI"

/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C11"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-C11 is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI).
TAG_TISSUE=RPE and Choroid
TAG_LIB=UI-E-C11
TAG_SEQ=ACCTA"

ORIGIN
Query Match 26.8%; Score 651.6; DB 5; Length 659;
Best Local Similarity 99.2%; Pred. No. 6.9e-113;
Matches 654; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1764 CGTCMAAGCTTCACAGGTGTCAAGAGCATAGACCAACAAAGCTTAAAGCTGTGA 1823
DB 659 CGTCMAAGCTTCACAGGTGTCAAGAGCATAGACCAACAAAGCTTAAAGCTGTGA 600
QY 1824 GTTCTGGGGCCCAAGAAAGTTTGAATTCCTCTAGAGAGGATGGGCTTATGAGGC 1883
DB 599 GTTCTGGGGCCCAAGAAAGTTTGAATTCCTCTAGAGAGGATGGGCTTATGAGGC 540
QY 1884 ACCGAGAAGTATCTCAAGTGAAGAGAAACTGTGAGTTTAACTGACGATATGCGCAG 1943
DB 539 ACCGAGAAGTATCTCAAGTGAAGAGAAACTGTGAGTTTAACTGACGATATGCGCAG 480
QY 1944 AGATCCCCGAAATACCTCAAGAAACCTTTGGAACATCAACCAACATACACACTA 2003
DB 479 AGATCCCCGAAATACCTCAAGAAACCTTTGGAACATCAACCAACATACACACTA 420
QY 2004 CACTCAAAAGATCAATGATCTTATTTGGGCTTGGAAAAACAGGATTAACACATTCT 2063
DB 419 CACTCAAAAGATCAATGATCTTATTTGGGCTTGGAAAAACAGGATTAACACATTCT 360
QY 2064 AACCTGCTTCTTATGGGGATGCTTGGCAGCCAGGTCCTACCTGTGTGTACACAGCA 2123
DB 359 AACCTGCTTCTTATGGGGATGCTTGGCAGCCAGGTCCTACCTGTGTGTACACAGCG 300
QY 2124 GGAACATGATACAGTCAACAGCCATACAGCTGTCCACACTGAAGAACGTGTCTCAACAG 2183
DB 299 GGAACATGATACAGTCAACAGCCATACAGCTGTCCACACTGAAGAACGTGTCTCAACAG 240
QY 2184 CTGAAATCAAAATGTTAGCTTAAATAGATAAATCCAGACTACTTCAAGCTTTAATGCC 2243
DB 239 CTGAAATCAAAATGTTAGCTTAAATAGATAAATCCAGACTACTTCAAGCTTTAATGCC 180
QY 2244 TTTTATTCATTAATAAAGTGAAGACTGACGACCAATGGAACATTTAATCAGATC 2303
DB 179 TTTTATTCATTAATAAAGTGAAGACTGACGACCAATGGAACATTTAATCAGATC 120
QY 2304 TGATTCAGAGTCCGGAAACCTTAGTCTATCTGATTCAGAGACCAAGCAGCCTTAGAT 2363
DB 119 TGATTCAGAGTCCGGAAACCTTAGTCTATCTGAATCAGAGACCAAGCAGCCTTAGAT 60
QY 2364 ACTGCCCAACTAATGATTTAATTAATCAATATCTGTTAAAAAAGCTTAAAAA 2422
DB 59 ACTGCCCAACTAATGATTTAATTAATCAATATCTGTTAAAAAAGCTTAAAAA 1

RESULT 4
 BU731809/c
 LOCUS
 DEFINITION UI-E-C11-*alt*-g-11-0-UI-s1 UI-E-C11 Homo sapiens cDNA clone
 ACCESSION BU731809
 VERSION BU731809.1 GI:23657073
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 674)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.reagen.com).
 Seq primer: M13 FORWARD
 Polya=yes.

FEATURES

Location/Qualifiers
 1..674
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-C11-*alt*-g-11-0-UI"
 /tissue_type="RPE and Choroid"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_1b="UI-E-C11"
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: Ecor I; Site_2: Not I; UI-E-C11 is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag (drr)18 tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
 TAG_TISSUE=RPE and Choroid
 TAG_LIB=UI-E-C11
 TAG_SEQ=ACCTA"

ORIGIN

Query Match 26.7%; Score 648.8; DB 5; Length 674;
 Best Local Similarity 97.9%; Pred. No. 2.8e-130;
 Matches 656; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 1751 GAACCATCAGCGCGCTCAAGCTTCAAGTGTACACAGGATAGACCAAGCAAAAGC 1810
 Db 670 GCACGAGAGCGCGCTCAAGCTTCAAGTGTACACAGGATAGACCAAGCAAAAGC 611

QY 1811 TTAAGACTGTAGTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGG 1870
 Db 610 TTTAAGACTGTAGTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGG 551
 QY 1871 GCGTTGATGAGACACCCAGAAAGTATCTCAAGTGTAGAGAAAGATGTGAGTTTAACTG 1390
 Db 550 GCGTTGATGAGACACCCAGAAAGTATCTCAAGTGTAGAGAAAGATGTGAGTTTAACTG 491
 QY 1931 ACGGATATGAGAGATCCCGGAAATATCACTCAAGAACCTTTGGAACATCAACACC 1990
 Db 490 ACGGATATGAGAGATCCCGGAAATATCACTCAAGAACCTTTGGAACATCAACACC 431
 QY 1991 AACATACACTTACACTCAAGATCATCATGATATCTTATTTGGGCTTTGGAACAGGAT 2050
 Db 430 AACATACACTTACACTCAAGATCATCATGATATCTTATTTGGGCTTTGGAACAGGAT 371
 QY 2051 GAAGCACTTCTTAACTGCTTCTTAAATGGGATGCTTGCCAGCAGGTCCTCACTGT 2110
 Db 370 GAAGCACTTCTTAACTGCTTCTTAAATGGGATGCTTGCCAGCAGGTCCTCACTGT 311
 QY 2111 GTGTACACAGCAGACACTGATCCAGTCAAGCCATACAGCTGTCCACATGAGAAG 2170
 Db 310 GTGTACACAGCAGCAGACACTGATCCAGTCAAGCCATACAGCTGTCCACATGAGAAG 251
 QY 2171 TGTCCTTCAACAGCCTGATCAATGATGTTAGCTTAATGATTAATCCAGACTTCTC 2230
 Db 250 TGTCCTTCAACAGCCTGATCAATGATGTTAGCTTAATGATTAATCCAGACTTCTC 191
 QY 2231 AGCCTTAAATGCTTTTAAATCAATTAATGAAAGCTGAGCTGACCATTTGAAACAT 2290
 Db 190 AGCCTTAAATGCTTTTAAATCAATTAATGAAAGCTGAGCTGACCATTTGAAACAT 131
 QY 2291 TTAAGTCAAGCTGTGATTCAGAGTGGGAAACCTTATGATTCATGATCCAAAGCAGC 2350
 Db 130 TTAAGTCAAGCTGTGATTCAGAGTGGGAAACCTTATGATTCATGATCCAAAGCAGC 71
 QY 2351 CACACCTTGTATATCTGCGCCCAATCAATGATTTAATTAATCAATCTGTTAAAAA 2410
 Db 70 CACACCTTGTATATCTGCGCCCAATCAATGATTTAATTAATCAATCTGTTAAAAA 11
 QY 2411 AAAAAAAAAA 2420
 Db 10 AAAAAAAAAA 1

RESULT 5
 LOCUS BI480798 666 bp mRNA linear EST 28-FEB-2002
 DEFINITION H2RPE-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA
 sequence.
 ACCESSION BI480798
 VERSION BI480798.1 GI:18998607
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 666)
 AUTHORS Buraczynska,M., Wears,A.V., Zarepari,S., Farjo,R., Filipova,E., Yuan,Y., MacVee,S.P., Hughes,B. and Swaroop,A.
 TITLE Towards an expression profile of native human retinal pigment epithelium: Identification of a non-redundant set of more than 1100 genes

JOURNAL Unpublished (2001)
 COMMENT Contact: Swaroop, A.
 Department of Ophthalmology and Visual Sciences
 Kellogg Eye Center, University of Michigan
 540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
 Tel: 734 615 2246
 Fax: 734 647 0228
 Email: swaroop@umich.edu

PCR Primers
 FORWARD: M13/PUC-Reverse - ccagcagcagcgtgtaaacg
 BACKWARD: M13/PUC-Forward - agcgataacaattccacacag
 Seq primer: M13/PUC-Reverse.

FEATURES
 SOURCE
 Location/Qualifiers
 1. .666
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="Native" Retinal Pigment Epithelium sheets"
 /dev_stage="juvenile"
 /clone_lib="Human Retinal Pigment Epithelium (2)"
 /note="Organ: Retina; Vector: pSport1"

ORIGIN
 Query Match 26.4%; Score 640.4; DB 4; Length 666;
 Best Local Similarity 99.4%; Pred. No. 1.9e-128;
 Matches 652; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

10 CCACCAAGCTAGTGGCCAGACCTTCTGTGGATCATCGGACCCACCTGAACCCACCTG 69
12 CCACCAAGCTAGTGGCCAGACCTTCTGTGGATCATCGGACCCACCTGAACCCACCTG 71
70 ACCCAAGCCCACTGCTGACAGCCCACTGCTGGCCATGACATCACTTACAAAGCCAA 129
72 ACCCAAGCCCACTGCTGACAGCCCACTGCTGGCCATGACATCACTTACAAAGCCAA 131
130 TGGCTAATGCGGCTTGAAGCTCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 189
132 TGGCTAATGCGGCTTGAAGCTCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 191
190 ACAAGCTGCTATATGAGGAGTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTT 249
192 ACAAGCTGCTATATGAGGAGTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTT 251
250 TTATAGGCTGAGCCCTCAGGAGAAACAAGCTGATGTTTGAAGAACTGAATCTGAT 309
252 TTATAGGCTGAGCCCTCAGGAGAAACAAGCTGATGTTTGAAGAACTGAATCTGAT 311
310 GCGACAGTATATCAAGTCTATCCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 369
312 GCGACAGTATATCAAGTCTATCCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCT 371
370 TCGGAGCCCGCTGAGGAAACAGTACGAACTGCGGTGAGCCGCTGCTGCTGCTGCT 429
372 TCGGAGCCCGCTGAGGAAACAGTACGAACTGCGGTGAGCCGCTGCTGCTGCTGCT 431
430 TGGTGTGGGCTTGTGCAAGGCAAGGAGCAAGGCGGCTGCTGCTGCTGCTGCTGCT 489
432 TGGTGTGGGCTTGTGCAAGGCAAGGAGCAAGGCGGCTGCTGCTGCTGCTGCTGCT 491
490 TCCGCTAAGCCCACTGAGGCAAGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCT 549
492 TCCGCTAAGCCCACTGAGGCAAGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCT 551
550 AGGCGTTCCTCCAGGCGCCAGCACTGCTGCAAGGAGGCTTATGATCTCCGCAAGCA 609
552 AGGCGTTCCTCCAGGCGCCAGCACTGCTGCAAGGAGGCTTATGATCTCCGCAAGCA 610
610 AGCAGTTGAGAAACTGAGCTTACACACATGTTCTGGAGTCCCTGAGGTGTG 665
611 AGCAGTTGAGAAACTGAGCTTACACACATGTTCTGGAGTCCCTGAGGTGTG 666

```

RESULT 6
 CA389968 639 bp mRNA linear EST 06-NOV-2002
 LOCUS ca104n03.Y1 Human Retinal pigment epithelium/choroid cDNA
 DEFINITION (Un-normalized, unamplified): cs Homo sapiens cDNA clone ca104n03
 5', mRNA sequence.
 ACCESSION CA389968
 VERSION CA389968.1 GI:24720628
 KEYWORDS EST.

SOURCE
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Wistow, G., Bernstein, S.L., Wyatt, M.K., Farris, R.N., Behal, A.,
 Touchman, J.W., Bouffard, G., Smith, D., and Peterson, K.
 Expressed sequence tag analysis of human RPE/choroid for the
 NEBank Project: Over 6000 non-redundant transcripts, novel genes
 and splice variants
 Mol. Vis. 8 (4), 205-220 (2002)
 MEDLINE 22103460
 PUBMED 12107410

COMMENT
 Contact: Wistow G
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: graeme@helix.nih.gov
 Plate: 104 row: h column: 03
 Seq primer: M13Rpl reverse primer (ABI).

FEATRES

source
 Location/Qualifiers

```

1. .639
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="ca104n03"
/tissue_type="RPE/choroid"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_lib="Human Retinal pigment epithelium/choroid cDNA
(Un-normalized, unamplified): cs"
/note="Organ: Eye; Vector: pCMVSPORTs; Two different donor
eyes (75-80 years old) yielded approximately 600 mg of
dissected RPE/choroid tissue. This in turn yielded 340 ug
of total RNA and 7 ug of mRNA. A directionally cloned cDNA
library in the pCMVSPORTs vector was constructed at Life
Technologies (Rockville, MD; now part of Invitrogen Corp),
essentially following the protocols of the Superscript
Plasmid System (Invitrogen Corp).
<http://www.invitrogen.com/>". The library code
designation was cs. For this library, cDNA inserts were
cloned into the NotI/MluI sites of the vector. EST
analysis was performed on the unamplified library at the
NIH Intramural Sequencing Center (NISC)."
```

ORIGIN

Query Match 26.3%; Score 638; DB 6; Length 639;
 Best Local Similarity 99.8%; Pred. No. 6.3e-128;
 Matches 638; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

12 ACCAGCTAGTGGCAGACCTTCTGTGGATCATGGAACCACTGAAACCCACCTGAC 71
1 ACCAGCTAGTGGCAGACCTTCTGTGGATCATGGAACCACTGAAACCCACCTGAC 60
72 CCAAGCCCACTGCTGACAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 131
61 CCAAGCCCACTGCTGACAGCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
132 GCTAATGCGGCTTGAAGCTCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 191
121 GCTAATGCGGCTTGAAGCTCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
192 AAGCTGCTATATGAGGAGTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTTAT 251
181 AAGCTGCTATATGAGGAGTTCTTAATCTTCTGCTGCTGCTGCTGCTGCTGCTTAT 240
252 TATAGGCTGAGCCCTCAGGAGAAACAAGCTGATGTTTGAAGAACTGAATCTGAT 311
241 TATAGGCTGAGCCCTCAGGAGAAACAAGCTGATGTTTGAAGAACTGAATCTGAT 300
312 GACAGTATATCCAGCTCATCCCATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 371

```

```

Db      301 GACAGCTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCTACGTGAGCGCTGTC
Qy      372 GTGACCCCGCTGTGGAAACAGTACGAGAACTCGCTGGCCCGACCGCTCATGAGCTG
Db      361 GTGACCCCGCTGTGGAAACAGTACGAGAACTCGCTGGCCCGACCGCTCATGAGCTG
Qy      432 GTGTCGGGCTTCTGTCGAAAGCAAGAGCAAGCGGCGCTGCGCGGCGGCGGCGGCTGTC
Db      421 GTGTCGGGCTTCTGTCGAAAGCAAGAGCAAGCGGCGCTGCGCGGCGGCGGCTGTC
Qy      492 CGTACGCCCAACCTGGGCAACGTGCTATCTCTGCGAGCGCTCAGACCGGACGTCTCAAG
Db      481 CGTACGCCCAACCTGGGCAACGTGCTATCTCTGCGAGCGCTCAGACCGGACGTCTCAAG
Qy      552 CGCTTCCCGACCGCGCCAGCACTGTGATCTTATGACTCCGGAGAAACAAG
Db      541 CGCTTCCCGACCGCGCCAGCACTGTGATCTTATGACTCCGGAGAAACAAG
Qy      612 CAGTTGAGAAACTGAGCTTACCAACACATGTTCTGG
Db      601 CAGTTGAGAAACTGAGCTTACCAACACATGTTCTGG

```

RESULT 7
CA397981 636 bp mRNA linear EST 06-NOV-2002
LOCUS CA397981
DEFINITION cs98c05.y1 Human Retinal pigment epithelium/choroid cDNA
(Un-normalized, unamplified): cs Homo sapiens cDNA clone cs98c05
5', mRNA sequence.

ACCESSION CA397981
VERSION CA397981
KEYWORDS EST.
SOURCE GI:24735789

ORGANISM Homo sapiens (human)
Homo sapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 636)

TITLE Wistow G., Bernstein, S.L., Wyatt, M.K., Pariss, R.N., Behal, A.,
Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.
Expressed sequence tag analysis of human RPE/choroid for the
Neibank Project: Over 6000 non-redundant transcripts, novel genes
and splice variants

JOURNAL Mol. Vis. 8 (4), 205-220 (2002)
MEDLINE 22103460
PUBMED 12107410

COMMENT

Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: gireme@helix.nih.gov
Plate: 98 row: C column: 05
Seq primer: M13RP1 reverse primer (ABI).
Location/Qualifiers
1..636

FEATURES
Source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cs98c05"
/tissue_type="RPE/choroid"
/dev_stage="Adult"
/lab_host="EMD10B"
/clone_id="Human Retinal pigment epithelium/choroid cDNA
(Un-normalized, unamplified): cs
/note="Organ: Eye; Vector: pCMVSPORT6; Two different donor
eyes (75-80 years old) yielded approximately 600 mg of
dissected RPE/choroid tissue. This in turn yielded 340 ug
of total RNA and 7 ug of mRNA. A directionally cloned cDNA
library in the pCMVSPORT6 vector was constructed at Life
Technologies (Rockville, MD; now part of Invitrogen Corp),
essentially following the protocols of the Superscript

ORIGIN

Query Match 26.0%; Score 631.2; DB 6; Length 636;
Best Local Similarity 99.5%; Pred. No. 1.9e-126;
Matches 633; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Plasmid System (Invitrogen Corp.
<http://www.invitrogen.com/>) The library code
designation was cs. For this library, cDNA inserts were
cloned into the NotI/Mut sites of the vector. EST
analysis was performed on the unamplified library at the
NIH Intramural Sequencing Center (NISC)."

```

Qy      1328 CCCGAGCCACAGCCCCCTTACACAGCTGCTTCGCCAGTTCCGAGGCTCTTATG
Db      1 CCCGAGCCACAGCCCCCTTACACAGCTGCTTCGCCAGTTCCGAGGCTCTTATG
Qy      1388 GGCCTCACTTCACATCAGCTGAAACAAAGAGATGAGATTCCAGCCCAATCGAG
Db      61 GGCCTCACTTCACATCAGCTGAAACAAAGAGATGAGATTCCAGCCCAATCGAG
Qy      1448 GACGAGGAGATGCTACGCTGGCATCATTTGCGCTTCTAGAGCTGAGTCCATGAT
Db      121 GACGAGGAGATGCTACGCTGGCATCATTTGCGCTTCTAGAGCTGAGTCCATGAT
Qy      1508 CACCATCTCCCAAGGGCAAACTCAGAGCCAAACTATCTGTGCCCAGAGGGAATCCCTT
Db      181 CACCATCTCCCAAGGGCAAACTCAGAGCCAAACTATCTGTGCCCAGAGGGAATCCCTT
Qy      1568 CTCACGAGGGCTGCGCCCAAAAACACAGAGGCGCAACAGAGTTAGGGGCCAGAGAA
Db      241 CTCACGAGGGCTGCGCCCAAAAACACAGAGGCGCAACAGAGTTAGGGGCCAGAGAA
Qy      1628 GACAAACAGGCTCTGGAAGCTTAAAGCTGTGACGCTTCAAGTCTGGCCACTGTATGAG
Db      301 GACAAACAGGCTCTGGAAGCTTAAAGCTGTGACGCTTCAAGTCTGGCCACTGTATGAG
Qy      1688 AGGCCAGGCTACTACAGTGCCTCCCAACAGAGCGCCCTCAGCCCATGTTCTTCCC
Db      361 AGGCCAGGCTACTACAGTGCCTCCCAACAGAGCGCCCTCAGCCCATGTTCTTCCC
Qy      1748 CTGAAACATTCAGCGCGCTGAAGCTTCAACAGTGTCAAGGCAATGAGACCAAGACAA
Db      421 CTGAAACATTCAGCGCGCTGAAGCTTCAACAGTGTCAAGGCAATGAGACCAAGACAA
Qy      1808 AGCTTAAAGACGTGAGTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGCGAT
Db      481 AGCTTAAAGACGTGAGTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGCGAT
Qy      1868 GGGGCTTGTATGAGACACCCAGAGATCTCAAGTGAAGAGAAACTGTGAGTTTAA
Db      541 GGGGCTTGTATGAGACACCCAGAGATCTCAAGTGAAGAGAAACTGTGAGTTTAA
Qy      1928 CTGACGATATGCGCAGAGATCCCGAAATCACTC
Db      601 CTGACGATATGCGCAGAGATCCCGAAATCACTC

```

RESULT 8
BU731149/c 651 bp mRNA linear EST 09-OCT-2002
LOCUS BU731149
DEFINITION UT-E-C11-afg-j-20-0-UT.s1 UT-E-C11 Homo sapiens cDNA clone
BU731149
ACCESSION BU731149
VERSION BU731149.1 GI:23655752
KEYWORDS EST.
SOURCE Homo sapiens (human)
Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 651)
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com).
Seq primer: M13 FORWARD
POLY-A=yes.

FEATURES

source
1..651
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C11-afq-j-20-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C11"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-C11 is a normalized cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG_TISSUE=RPE and Choroid
TAG_LIB=UI-E-C11
TAG_SEQ=ACCTA"

ORIGIN

Query Match 25.8%; Score 626.6; DB 5; Length 651;
Best Local Similarity 99.2%; Pred. No. 1.9e-125;
Matches 629; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY ACAAGGCGCCAAACAGAACTTGAAGGGCCAGAGAGACAAACAGGCTTGAAGTTAAG 1652
DB 651 ACAAGGCGCCAAACAGAACTTGAAGGGCCAGAGAGACAAACAGGCTTGAAGTTAAG 592
QY CTGTGAGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGCTACTTACAGTGC 1712
DB 591 CTGTGAGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGCTACTTACAGTGC 532
QY AGAGGCGCCCTCAGGCGCCCACTCCAGTCTTCCCTTGAACCAACAGGCGCTCAAGC 1772
DB 531 AGAGGCGCCCTCAGGCGCCCACTCCAGTCTTCCCTTGAACCAACAGGCGCTCAAGC 472
QY TTTCACAGTGTACAGGCGCATAGACCAAAAGCTTAAAGACTGTGAGTTCTGGGG 1832
DB 471 TTTCACAGTGTACAGGCGCATAGACCAAAAGCTTAAAGACTGTGAGTTCTGGGG 412
QY CCAAGAAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGTATGAGAGACCCAGAG 1892
DB 411 CCAAGAAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGTATGAGAGACCCAGAG 352

QY 1893 TATCTCAAGTAGAGAGAAAACCTGTGAGTTAACTTGACCGGATATGCCAGATCCCG 1952
DB 351 TATCTCAAGTAGAGAGAAAACCTGTGAGTTAACTTGACCGGATATGCCAGATCCCG 292
QY 1953 AAAATCACCTTAAAGAACTTTGGACATATCAACCAACATATACACTACACTCAAG 2012
DB 291 AAAATCACCTTAAAGAACTTTGGACATATCAACCAACATATACACTACACTCAAG 232
QY 2013 ATTCATGAGATCCTTATTTGGGCTTTGGAAAACAGGATGAAGACATTCCTTACCTGCTT 2072
DB 231 ATTCATGAGATCCTTATTTGGGCTTTGGAAAACAGGATGAAGACATTCCTTACCTGCTT 172
QY 2073 CCTATGAGGATGCTTGGCCAGCCAGGCTCTTACCTGTGTGTAGACACGACGACATGA 2132
DB 171 CCTATGAGGATGCTTGGCCAGCCAGGCTCTTACCTGTGTGTAGACACGACGACATGA 112
QY 2133 TCACATCAGCCATATACGCTGTCCACATGAAAGACCTGTCTTACACAGCTGATCA 2192
DB 111 TCACATCAGCCATATACGCTGTCCACATGAAAGACCTGTCTTACACAGCTGATCA 52
QY 2193 AATGTTAGCTTATATGATTAATAATCCAGACTA 2226
DB 51 AATGTTAGCTTATATGATTAATAATCCAGACTA 18

RESULT 9

BM663028/c 592 bp mRNA linear EST 27-FEB-2002
LOCUS
DEFINITION
UI-E-C10-aad-h-10-0-UI.s1 UI-E-C10 Homo sapiens cDNA clone
BM663028
ACCESSION
VERSION
KEYWORDS
BM663028.1 GI:18968017
EST.
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
AUTHORS
TITLE
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL
MEDLINE
PUBMED
COMMENT
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548

Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com).
Seq primer: M13 Forward
POLY-A=yes.

FEATURES

source
1..592
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C10-aad-h-10-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C10"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-C10 is a cDNA library containing the following tissue(s): RPE and Choroid. The library was constructed

according to Bonaldo, Lemmon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACCCTA. This library was created for the program, Gene Discovery in the Vastul System, supported by National Eye Institute (NEI).
TAG_TISSUE=RPE and Choroid
TAG_LIB=UI-E-C10
TAG_SEQ=ACCCTA"

ORIGIN

Query Match 24.0%; Score 582.4; DB 4; Length 592;
Best Local Similarity 99.0%; Pred. No. 7.8e-116;
Matches 586; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1830 GGGCCAGAAAAGTTTGAATTCCTCAGAGAGCGAGGGCTTGATGAGACCCAG 1889
DB 592 GGGCCAGAAAAGTTTGAATTCCTCAGAGAGGTATGGGCTTGATGAGACCCAG 533
QY 1890 AAGTATCTCAAGTGAAGAGAAAACCTGTGAGTTTAACTGAAGATATGCCAGAGATCC 1949
DB 532 AAGTATCTCAAGTGAAGAGAAAACCTGTGAGTTTAACTGAAGATATGCCAGAGATCC 473
QY 1950 CCGAAATACACCTCAAGAAACCTTTGGAACATACCCAAACATACCACTACCTCA 2009
DB 472 CCGAAATACACCTCAAGAAACCTTTGGAACATACCCAAACATACCACTACCTCA 413
QY 2010 AAGATCAATGATCTTATTTGGCCCTTGGAAAACAGGGATAGACATTCCTAACCTG 2069
DB 412 AAGATCAATGATCTTATTTGGCCCTTGGAAAACAGGGATAGACATTCCTAACCTG 353
QY 2070 CTTCCTCAATGGGGAGTCTTGCCAGACAGGTCTCACTGTGTGACACAGCAGAGAC 2129
DB 352 CTTCCTCAATGGGGAGTCTTGCCAGACAGGTCTCACTGTGTGACACAGCAGAGAC 293
QY 2130 TGATCAGTCAAGCATACAGCTGTCCACACTGAAAGACGTGTCTCAACAGCTGAA 2189
DB 292 TGATCAGTCAAGCATACAGCTGTCCACACTGAAAGACGTGTCTCAACAGCTGAA 233
QY 2190 TCAATGTGTACCTTAATGATTAATAATCCAGACTTCAAGCTTTAATGCTTTTAT 2249
DB 232 TCAATGTGTACCTTAATGATTAATAATCCAGACTTCAAGCTTTAATGCTTTTAT 173
QY 2250 TCATAAAAACTGTGAAGCTAGACGAAACATTTGAAACATTTAATCTGAGCTGAT 2309
DB 172 TCATAAAAACTGTGAAGCTAGACGAAACATTTGAAACATTTAATCTGAGCTGAT 113
QY 2310 CAGAGTGGGAAACCTTGAATTTCTATCTGAATCCAGACGACACCTTATATATGCTCC 2369
DB 112 CAGAGTGGGAAACCTTGAATTTCTATCTGAATCCAGACGACACCTTATATATGCTCC 53
QY 2370 CAATATATGATTTAATAATCAAAATCTGTTAAAAAATAAAAA 2421
DB 52 CAATATATGATTTAATAATCAAAATCTGTTAAAAAATAAAAA 1

RESULT 10
BE410951 773 bp mRNA linear EST 21-JUL-2000
LOCUS 6013036621 NIH MGC 21 Homo sapiens cDNA clone IMAGE:3638175 5',
DEFINITION mRNA sequence.
ACCESSION BE410951
VERSION BE410951.1 GI:9347401
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 773)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LNCM38 row: 1 column: 16
High quality sequence stop: 662.

FEATURES

source

1. .773
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3638175"
/issue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 21"
/notes="Organ: placenta; Vector: pORF7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGG(6). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 23.8%; Score 577; DB 2; Length 773;
Best Local Similarity 99.0%; Pred. No. 1.2e-114;
Matches 591; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1456 GGATGTCAAGCTGGGATGATGGCGGCTTCTAGGCGCGAGTCCCATGATCAACATCC 1515
DB 1 GGATGTCAAGCTGGGATGATGGCGGCTTCTAGGCGCGAGTCCCATGATCAACATCC 60
QY 1516 TCCAGAGGCAAACTCAAGAGCAAACTACTGTGAGCCCAAGAGGAATCCCTTCCACGA 1575
DB 61 TCCAGAGGCAAACTCAAGAGCAAACTACTGTGAGCCCAAGAGGAATCCCTTCCACGA 120
QY 1576 GGGCGTGGCC-AAAAACCAAGGCAAGCCCAACAAAGATTAGGGCCCAAGAGCAACA 1634
DB 121 GGGCGTGGCCCAAAACCAAGGCAAGCCCAACAAAGATTAGGGCCCAAGAGCAACA 180
QY 1635 AGGCTGTGAAGCTTAAGGTGTGACGCTTCAAGCTGTGACCACTGATCAGAGGCGAG 1694
DB 181 AGGCTGTGAAGCTTAAGGTGTGACGCTTCAAGCTGTGACCACTGATCAGAGGCGAG 240
QY 1695 GCTACTACAGTGGCCCAAGAGAGGCGCTTCAAGCTGTGACCACTGATCAGAGGCGAG 1754
DB 241 GCTACTACAGTGGCCCAAGAGAGGCGCTTCAAGCTGTGACCACTGATCAGAGGCGAG 300
QY 1755 CATCAGCGCGCTCAAGAGCTTCAAGTGTGACAGGATGAGACCAAGAGCAAAAGCTTAA 1814
DB 301 CATCAGCGCGCTCAAGAGCTTCAAGTGTGACAGGATGAGACCAAGAGCAAAAGCTTAA 360
QY 1815 AGACTGTGATTTGGGGCCAAAGAAAGTTTGAATTCCTCAGAGAGCGATGGGCGCT 1874
DB 361 AGACTGTGATTTGGGGCCAAAGAAAGTTTGAATTCCTCAGAGAGCGATGGGCGCT 420
QY 1875 TGATGAGACACCCAGAAATATCTCAAGTGAAGAGAAAACCTGTGAGTTTAACTGACGG 1934
DB 421 TGATGAGACACCCAGAAATATCTCAAGTGAAGAGAAAACCTGTGAGTTTAACTGACGG 480
QY 1935 ATATGCCAGAGATCCCGAAATATCACTCAAGAAACCTTTGGAACATTCACCAACACA 1994
DB 481 ATATGCCAGAGATCCCGAAATATCACTCAAGAAACCTTTGGAACATTCACCAACACA 540

QY 1995 TACACACTACACTCAAGATCATGATCTTATTGGCCCTTGGAAAAAGGATG 2051
|||||
Db 541 TACACACTACACTCAAGATCATGATCTTATTGGCCCTTGGAAAAAGGATG 597
|||||

RESULT 11
BM707948 585 bp mRNA linear EST 28-FEB-2002
LOCUS UT-E-C11-8-ft-g-11-0-UI r1 UI-E-C11 Homo sapiens cDNA clone
DEFINITION UT-E-C11-8-ft-g-11-0-UI 5', mRNA sequence.
ACCESSION BM707948
VERSION BM707948.1 GI:19021206
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 585)
AUTHORS Bernaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704447
PubMed 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.reagen.com).
Seq primer: M13 Reverse.
FEATURES
source
1..585
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-C11-8-ft-g-11-0-UI"
/tissue_type="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-C11"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-C11 is a normalized cDNA library containing the
following tissue(s) : RPE and Choroid. The library was
constructed according to Bernaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."

ORIGIN
Query Match 23.7%; Score 575; DB 4; Length 585;
Best Local Similarity 98.8%; Pred No. 3.2e-114;
Matches 578; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1759 AGCGCGTCAAGCTTCAAGTGTACAGGATAGACCAAAAGCTTTAAAGAC 1818
|||||
Db 1 AGCGCGTCAAGCTTCAAGTGTACAGGATAGACCAAAAGCTTTAAAGAC 60
|||||

QY 1819 TGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCTTGAT 1878
|||||
Db 61 TGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGTATGGGGCTTGAT 120
|||||

QY 1879 GGAGCACCAGAGATATCTCAAGTGAAGAGAAAACTGTGAGTTTAACTGACGGATAT 1938
|||||
Db 121 GGAGCACCAGAGATATCTCAAGTGAAGAGAAAACTGTGAGTTTAACTGACGGATAT 180
|||||

QY 1939 GCCAGAGATCCCCGAAAAATCACTCAAGAATCTTTGGAACAATCAACCAACATCA 1998
|||||
Db 181 GCCAGAGATCCCCGAAAAATCACTCAAGAATCTTTGGAACAATCAACCAACATCA 240
|||||

QY 1999 CACTACACTCAAGATCATGATGCTTATTGGGCTTGGAAAAAGGATGAAGCACA 2058
|||||
Db 241 CACTACACTCAAGATCATGATGCTTATTGGGCTTGGAAAAAGGATGAAGCACA 300
|||||

QY 2059 TTCTTAACCTGCTTCTTAATGGGATGCTTGGCAAGCCAGTCTCACTGTGTACAC 2118
|||||
Db 301 TTCTTAACCTGCTTCTTAATGGGATGCTTGGCAAGCCAGTCTCACTGTGTACAC 360
|||||

QY 2119 CAGCAGACATGATCCAGTACAGCCATACAGCTGTCTCAACATGAAAGATGTCTAC 2178
|||||
Db 361 CAGCAGACATGATCCAGTACAGCCATACAGCTGTCTCAACATGAAAGATGTCTAC 420
|||||

QY 2179 AACAGCGTAATCAAGTGTAGCTTAATAGATAAAATCCAGACTACTTCAGGCTTTA 2238
|||||
Db 421 AACAGCGTAATCAAGTGTAGCTTAATAGATAAAATCCAGACTACTTCAGGCTTTA 480
|||||

QY 2239 ATGCGTTTATTCATTAATAAACTGTGAAGCTAGACTGAACCATTTGAACATTTAACTCA 2298
|||||
Db 481 ATGCGTTTATTCATTAATAAACTGTGAAGCTAGACTGAACCATTTGAACATTTAACTCA 540
|||||

QY 2299 GACTCTGATTCAGAGTCCGGAACCCCTTAAGTTCTATCTGATCCA 2343
|||||
Db 541 GACTCTGATTCAGAGTCCGGAACCCCTTAAGTTCTATCTGATCCA 585
|||||

RESULT 12
BE385296 593 bp mRNA linear EST 21-JUL-2000
LOCUS 60127572F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3618665 5',
DEFINITION mRNA sequence.
ACCESSION BE385296
VERSION BE385296.1 GI:9330661
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 593)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at: image.llnl.gov
Plate: L10M287 row: 0 column: 18
High quality sequence stop: 593.
FEATURES
source
1..593
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3618665"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_20"

/note="Organ: skin; Vector: pOT87; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

Query Match 23.3%; Score 565.2; DB 2; Length 593;
Best Local Similarity 99.3%; Pred. No. 4.3e-112;
Matches 578; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

1471 CATCATTTGGCGCTTCTAGAGCTGAGTCCCATGATCCATCTCCGAGGGCAAACTC 1530
1 CATCATTTGGCGCTTCTAGAGCTGAGTCCCATGATCCATCTCCGAGGGCAAACTC 60
1531 AAGGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTCCAGAGGGCTTCCC-AAAA 1589
61 AAGGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTCCAGAGGGCTTCCC-AAAA 120
1590 ACCAGAGGAGGCAAGCAAGAGAGTAAAGGGGCAAGAGAGAGAGAGAGAGAGAGAG 1649
121 ACCAGAGGAGGCAAGCAAGAGAGTAAAGGGGCAAGAGAGAGAGAGAGAGAGAGAG 180
1650 AAGCTGTGAGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCGAGGCTACTACAGTGC 1709
181 AAGCTGTGAGAGCGCTTCAAGTCTGGCCCACTGTATCAGAGGCGAGGCTACTACAGTGC 240
1710 CACAGAGCGCCCTCAGAGGCGGCACTCCCATGTTCTTCCCTTGAAGACCATCAGCGCTGAA 1769
241 CACAGAGCGCCCTCAGAGGCGGCACTCCCATGTTCTTCCCTTGAAGACCATCAGCGCTGAA 300
1770 AAGCTTCAAGTGTCAAGGAGGATAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1829
301 AAGCTTCAAGTGTCAAGGAGGATAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 360
1830 GGGCCCAAG 1889
361 GGGCCCAAG 420
1890 AAGTATCTCAAGTGTGAG 1949
421 AAGTATCTCAAGTGTGAG 480
1950 CCGAAATACCTCTCAAG 2009
481 CCGAAATACCTCTCAAG 540
2010 AAGATCATGATGATCTTATTTGGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2051
541 AAGATCATGATGATCTTATTTGGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 582

RESULT 13
BO879880 963 bp mRNA linear EST 16-AUG-2002
LOCUS BO879880
DEFINITION AGENCOURT 8241531 Lupski dorsal root ganglion Homo sapiens cDNA
AGENCOURT 8241531 5', mRNA sequence.
VERSION BO879880
KEYWORDS BO879880.1 GI:22271888
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 963)
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-1@mail.nih.gov

Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Consortium
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LNL13563 row: 1 column: 08
High quality sequence stop: 623.

FEATURES
source
Location/Qualifiers
1..963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski dorsal root ganglion"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGCTCCG-3' and
5'-GACTAGTCTAGATCGGAGGCGGCGGCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN

Query Match 23.2%; Score 564.6; DB 5; Length 963;
Best Local Similarity 77.3%; Pred. No. 5.8e-112;
Matches 795; Conservative 0; Mismatches 31; Indels 203; Gaps 1;

445 TCAGAGGCAAG 504
26 TCAGAGGCAAG 85
505 TGGGCAACGTGCTCATCTCTGCGAGGCTCAGACCGGAGTCTTACAGAGGCTTCCCAGCG 564
86 TGGGCAACGTGCTCATCTCTGCGAGGCTCAGACCGGAGTCTTACAGAGGCTTCCCAGCG 145
565 CCAGACCTGTGTGCAAG 624
146 CCAGACCTGTGTGCAAG 205
625 TGAAGCTTACCAACAATGTTCTGGGTGCTTGGGTGTTTCCAACTGTCAATGA 684
206 TGAAGCTTACCAACAATGTTCTGGGTGCTTGGGTGTTTCCAACTGTCAATGA 265
685 AAGCGTGTGTGAGAGTGAATCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 744
266 AAGCGTGTGTGAGAGTGAATCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 325
745 TGAACACCTTGGCTACTAGTGTGACACCTGTATGCTTACGATGATTTAGTATCCAC 804
326 TGAACACCTTGGCTACTAGTGTGACACCTGTATGCTTACGATGATTTAGTATCCAC 385
805 TGGTGTATACAGAGTGTGACTGTGGCGGTGTACAGCTTCTTCTGACTGTGTGTTG 864
386 TGGTGTATACAGAGTGTGACTGTGGCGGTGTACAGCTTCTTCTGACTGTGTGTTG 445
865 GCGGCAAGTTTGAACCAAG 924
446 GCGGCAAGTTTGAACCAAG 505
925 CCGTCTTACAGTCTCTGAGTGTCTTCTTATGTTGTGCTGAGTGAAGTGGGCTCTTCCA 984
506 CCGTCTTACAGTCTCTGAGTGTCTTCTTATGTTGTGCTGAGTGAAGTGGGCTCTTCCA 551
985 GGGCCCTGCTGGCTGAGAGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1044

Db 552 ----- 551
QY 1045 GGATGCACTGTGAGAAAGAAAGTCTACCGGTAGAAAGCAGCCAGGCTGTGGCCCA 1104
Db 552 ----- 551
QY 1105 CACCTGTAATCCAGACTACTCGGAGGCTGAGAGAGAAATCGCTGACCAGGAGGC 1164
Db 552 ----- 551
QY 1165 GGAGGTGTGTGTCAGAGCAGCTCATCAACCCCTTTGAGAGATGATGATTTTGA 1224
Db 552 ----- 602
QY 1225 GACCACTGATTTGTGTCAGAGAAATTTGTCAGGTGCTCTGTGGTGTGATGATGCA 1284
Db 603 GACCAACTGATTTGTGTCAGAGAAATTTGTCAGGTGCTCTGTGGTGTGATGATGCA 662
QY 1285 CCAGGACCTGCTCGATGAGAGCCGAGCATGTACTGAAATAAGCCGAGCCAGCCGCC 1344
Db 663 CCAGGACCTGCTCGATGAGAGCCGAGCATGTACTGAAATAAGCCGAGCCAGCCGCC 722
QY 1345 CTACACAGCTGCTTCCGCCAGTTCCGTGAGCCTCTTTATGAGCTCCACCTTCAACAT 1404
Db 723 CTACACAGCTGCTTCCGCCAGNTCCGTGAGCCTCTTTATGAGCTCCACCTTCAACAT 782
QY 1405 CAGCTGAAACAAGAGAGATGAGTTCAGCCCATAGAGAGACGAGAGATGCTCA 1464
Db 783 CAGCTGAAACAAGAGAGATGAGTTCAGCCCATAGAGAGACGAGAGATGCTCA 842
QY 1465 CGCTGGCAT 1473
Db 843 CTCGCCCT 851

RESULT 14
CD518675 881 bp mRNA linear EST 06-JUN-2003
LOCUS AGNCOURT_14375996 NIH_MGC_181 Homo sapiens cDNA clone
DEFINITION IMAGE:30397964 5', mRNA sequence.
ACCESSION CD518675
VERSION CD518675.1 GI:31450393
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS 1 (bases 1 to 881)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgsabde-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM482 row: b column: 21
High quality sequence scop: 652.
Location/Qualifiers
1..881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30397964"
/tissue_type="White Matter"

/dev_stage="Unknown"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_181"
/notes="Vector: pCMV-Sport6.1; Site 1: NotI; Site 2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."

Query Match 22.8%; Score 553.2; DB 6; Length 881;
Best Local Similarity 76.5%; Pred. No. 1.8e-109;
Matches 817; Conservative 0; Mismatches 43; Indels 208; Gaps 3;

QY 570 CACCTGGTGCAGAGAGGCTTATGATCCCGCAGAAACAGACAGTGGAGAACTGAGC 629
Db 1 CACCTGGTGCAGAGAGGCTTATGATCCCGCAGAAACAGACAGTGGAGAACTGAGC 60
QY 630 CTACACACACATGTTCTGGGTGCCCTGGGTGTGTTGCCAACCTGTCAATGAGGCG 689
Db 61 CTACACACACATGTTCTGGGTGCCCTGGGTGTGTTGCCAACCTGTCAATGAGGCG 120
QY 690 TGGCTTGAGGTGCAATTCGGGACCTTACTCTCTCCAGAGCTGTGCAAGATGCAAC 749
Db 121 TGGCTTGAGGTGCAATTCGGGACCTTACTCTCTCCAGAGCTGTGCAAGATGCAAC 180
QY 750 ACCTGGGTACTCAGTGTGACACTGTATGCTTACAGCTGAGATTATGATCCACTGCTG 809
Db 181 ACCTGGGTACTCAGTGTGACACTGTATGCTTACAGCTGAGATTATGATCCACTGCTG 240
QY 810 TATACACAGGTGTGACTGTGCGGTGTGACAGTCTTCTCTGACTGTGCTGATTGGGCG 869
Db 241 TATACACAGGTGTGACTGTGCGGTGTGACAGTCTTCTCTGACTGTGCTGATTGGGCG 300
QY 870 CAGTTTCTGAAACCAAGCCAGGCTTACCCTGGCCATGAGCTGACCTGTGTGCGCTC 929
Db 301 CAGTTTCTGAAACCAAGCCAGGCTTACCCTGGCCATGAGCTGACCTGTGTGCGCTC 360
QY 930 TTCAAGTTCGTCAGTCTTCTTCTATGTTGGCTGTGAGAGTGGGCTCTCCAGGCGC 989
Db 361 TTCAAGTTCGTCAGTCTTCTTCTATGTTGGCTGTGAGAGTGGGCTCTCCAGGCGC 401
QY 990 CTGCTGGCTGAGGAGCATGAGCCAGAGGGGTGATGCGCAGACTGCTTGAAGACGAGATG 1049
Db 402 ----- 401
QY 1050 CAGTGTGAGAAAGAAAGTCTCAAGGGTAGAAAGACAGCCAGCGTGTGCGCACACT 1109
Db 402 ----- 401
QY 1110 GTAATCCAGCTACTCGGAGAGCTGAGGAGAGAAATCGCTTGAACCCGGAGCGGAG 1169
Db 402 ----- 401
QY 1170 TTGTGTGTCAGAGCAGCTCATCAACCCCTTTGAGAGAGATGATGATTTTGAGACCA 1229
Db 402 ----- 457
QY 1230 ACTGATTTGTGACAGGAATTTGACAGTGTCCCTGTTGGCTGTGAGATGACACAGG 1289
Db 458 ACTGATTTGTGACAGGAATTTGACAGTGTCCCTGTTGGCTGTGAGATGACACAGG 517
QY 1290 ACCGTGCTCGATGAGAGCCGAGCATGTATCTGGAATTAAGCCCGAGACCAAGCCCCCTTCA 1349
Db 518 ACCGTGCTCGATGAGAGCCGAGCATGTATCTGGAATTAAGCCCGAGACCAAGCCCCCTTCA 577
QY 1350 CAGCTGCTTCCGCCAGTTCGTCGAGGCTCTTTATGAGGCTCCACCTTCAACATCAGCC 1409
Db 578 CAGCTGCTTCCGCCAGTTCGTCGAGGCTCTTTATGAGGCTCCACCTTCAACATCAGCC 637
QY 1410 TGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGACGAGAGAGATCTTCAAGCTG 1469
Db 638 TGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGACGAGAGAGATCTTCAAGCTG 697

QY 1470 GCATCATTTGGCGCTTCTCTAGGCTCTGAGTCCATGATCACCATCTCTCCAGGG--CAA 1527
 DB 698 GCATCATTTGGCGCTTCTCTAGGCTCTGAGTCCATGATCACCATCTCTCCAGGGCAAAC 757
 QY 1538 CTCAGAGCCAAACTACTGTGTGGCCCAAGAGGAA--TCCCTTCTCCAGAGGCTCTCC 1584
 DB 758 TCCAGGGACCAAACTACTGTGTGGCCCAAGAGGAAATTCCTTCTCCAGAGGCTCTGC 817
 QY 1585 CAAAAACCAAGCAGGCAACAGAGCTTAGGGGCGCAGGAAGCAA 1632
 DB 818 CCAAAAACCCCAAGCAGGCAACCAAGCTTTTAGGGGCGCCA 865
 RESULT 15
 LOCUS CO396179
 DEFINITION AGENCOURT_27872871 NIH_MGC_212 Homo sapiens cDNA clone
 IMAGE:30923539 5', mRNA sequence.
 ACCESSION CO396179
 VERSION CO396179.1 GI:49578095
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 756)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: c396b5-remail.nih.gov
 Tissue Procurement: Mary Hendrix
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: NDAMI164 row: 1 column: 20
 High quality sequence stop: 571.
 Location/Qualifiers
 1..756
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30923539"
 /issue_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH_MGC_212"
 /note="Organ: Lung; Vector: pYX-Abc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with Not I and then cloned
 directionally into pYX-Abc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GATAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN

Query Match 22.7%; Score 551.6; DB 7; Length 756;
 Best Local Similarity 99.1%; Pred. No. 3.9e-109;
 Matches 554; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 417 CGGCTCATGAGCGCTGTGTGGGCTGTGCAAGCAAGCAAGCAAGCGCGGCTGTG 476
 DB 8 CGCTCATGAGCGCTGTGTGGGCTGTGCAAGCAAGCAAGCAAGCGCGGCTGTG 67

QY 477 CGGCGCAGCTCATCCGCTACGCCAAGCTGAGCAACGCTGCTCATCTGCGCAGGCTGAGC 536
 DB 68 CGGCGCAGCTCATCCGCTACGCCAAGCTGAGCAACGCTGCTCATCTGCGCAGGCTGAGC 127
 QY 537 ACCGCACTTCAAGAGCGCTTCCCGAGGCCCAAGACCTGTGTGCAAGAGGCTTTATAGT 596
 DB 128 ACCGCACTTCAAGAGCGCTTCCCGAGGCCCAAGACCTGTGTGCAAGAGGCTTTATAGT 187
 QY 597 CCGGCAAGCAAGCAGTGTGGAAGAACTAGAGCTTACCAACAATGTTCTGGGCTGCC 656
 DB 188 CCGGCAAGCAAGCAGTGTGGAAGAACTAGAGCTTACCAACAATGTTCTGGGCTGCC 247
 QY 657 TGGGTGTGTTTGGCAACTGTCAATGAAAGCGTGTGAGGTGCAATCCGGACCTT 716
 DB 248 TGGGTGTGTTTGGCAACTGTCAATGAAAGCGTGTGAGGTGCAATCCGGACCTT 307
 QY 717 ATCTGCTCCAGAGCTGCTGGAACGAGATGAACACTTGGCTACTCAGTGTGACACCTG 776
 DB 308 ATCTGCTCCAGAGCTGCTGGAACGAGATGAACACTTGGCTACTCAGTGTGACACCTG 367
 QY 777 TATGCTACGACTGATTAATCCACTGTGTATACAGAGGTGATGATGAGCGGTG 836
 DB 368 TATGCTACGACTGATTAATCCACTGTGTATACAGAGGTGATGATGAGCGGTG 427
 QY 837 TACAGCTTCTTCTGACTGTGTAGTGTGGCGGAGTTTCTGAACCCAGCAAGGCTTAC 896
 DB 428 TACAGCTTCTTCTGACTGTGTAGTGTGGCGGAGTTTCTGAACCCAGCAAGGCTTAC 487
 QY 897 CTTGGCATAGCTGAGCTCTGTGTGCCCCGTCTTACGTTCTGTGAGTTCTTTAT 956
 DB 488 CTTGGCATAGCTGAGCTCTGTGTGCCCCGTCTTACGTTCTGTGAGTTCTTTAT 547
 QY 957 GTTGGCTGTGTAAGGTGG 975
 DB 548 GTTGGCTGTGTAAGGTGG 566

Search completed: March 26, 2005, 13:32:08
 Job time : 7816.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 21:57:50 ; Search time 1259.21 seconds
(without alignments)
11419.125 Million cell updates/sec

Title: US-09-622-964A-4
Perfect score: 2429
Sequence: 1 caggagagtcaccaccagccta.....aaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2429	100.0	2429	2	AAZ21228 Human CG1
2	2399.2	98.8	2404	12	ADQ84435 Human tum
3	2399.2	98.8	2404	13	ADQ83262 Human tum
4	2013	82.9	2229	2	AAZ21227 Human CG1
5	1541.8	63.5	1758	8	ABZ80972 Human CG1
6	1244.2	51.2	1263	2	AAV99722 Human adu
7	1008.2	41.5	1238	9	ADA44960 Human pol
8	658.2	27.1	1916	2	AAZ21229 Mouse CG1
9	642.2	26.4	16125	2	AAZ21226 Human CG1
10	642.2	26.4	16550	5	ABA14559 Human ner
11	642.2	26.4	18530	5	ABA14557 Human ner
12	640.6	26.4	18537	5	ABA14558 Human ner
13	587.8	24.2	1717	6	AB189697 Human pol
14	476.4	19.6	7108	5	ABA14556 Human ner
15	431.4	17.8	2112	6	ABQ61177 FLJ20132
16	431.4	17.8	2137	8	ABZ80973 Human bes
17	431	17.7	1530	8	ABZ80997 Human bes
18	350.4	14.4	1326	8	ACA03889 CDNA upre
19	341.4	14.1	1948	9	ACC59900 Human REM
20	339.4	14.0	1422	8	ABZ80975 Human bes

21	338.2	13.9	2500	11	ADM02614 Human CDN
22	336.8	13.9	2028	8	ABZ80974 Human bes
23	294.6	12.1	305	5	ABA11094 Human ner
24	243.4	10.0	630	13	ADQ53985 Novel can
25	237	9.8	2861	4	AB110793 Drosoph11
26	234.8	9.7	424	2	AAV86976 EST clone
27	219.2	9.0	1592	10	ADC29982 Human nov
28	213	8.8	620	5	ADL45614 Human ova
29	193	7.9	1608	4	AB112609 Drosoph11
30	189.2	7.8	1292	11	ACN89920 Breast ca
31	188.2	7.7	10760	4	AB110792 Drosoph11
32	175	7.2	1119	8	ABZ71972 Human ter
33	175	7.2	1198	6	ABK83853 Human CDN
34	175	7.2	1198	6	ABN95676 Gene #217
35	175	7.2	1234	5	AA67682 DNA encod
36	175	7.2	1279	5	AA85655 DNA encod
37	175	7.2	2183	5	AA891587 DNA encod
38	154.8	6.4	1526	5	AA872796 DNA encod
39	150.4	6.2	3592	4	AB112636 Drosoph11
40	146.8	6.0	1345	4	AB112637 Drosoph11
41	135.2	5.6	539	12	ACH74388 Human gen
42	134.2	5.5	235	12	ACH8088 Human sec
43	126.8	5.2	1350	4	AA76848 Human sec
44	124.4	5.1	2914	12	ADQ24621 Human sof
45	124.4	5.1	2914	12	ADQ25483 Human sof

ALIGNMENTS

RESULT 1	
AAZ21228	
ID	AAZ21228 standard; CDNA; 2429 BP.
XX	
AC	AAZ21228;
DT	22-NOV-1999 (first entry)
XX	
DE	Human CG1CE long form CDNA sequence.
XX	
KW	CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
KW	age-related macular dystrophy; ss.
XX	
OS	Homo sapiens.
XX	
FX	
FT	Key
FT	CDS
FT	Location/Qualifiers
XX	105..1412
XX	/*tag= a
XX	/product= "CG1CE short form protein"
XX	
XX	WO9943695-A1.
PD	
PD	02-SEP-1999.
XX	
PF	22-FEB-1999; 99WO-US003790.
XX	
PR	25-FEB-1998; 98US-0075941P.
PR	18-DEC-1998; 98US-0112926P.
XX	
PA	(MERI) MERCK & CO INC.
PA	(UYUP-) UNIV UPPSALA.
XX	
PI	Petrukhin K, Gaeky CT, Metzker M, Wadelius C;
XX	
DR	WPI; 1999-540560/45.
XX	
DR	P-PSDB; AAY29954.
XX	
PT	Human and mouse polynucleotides encoding CG1CE polypeptides.
XX	
PS	Claim 2; Fig 4; 67pp; English.
XX	
CC	The present sequence represents the human CG1CE CDNA sequence, which when
CC	mutated is responsible for Best's macular dystrophy (BMD).

CC Polynucleotides encoding GC1C2 are useful for diagnosing whether a
CC patient carries a mutation in the GC1C2 gene. Normal and mutated GC1C2
CC proteins are useful for identifying activators and/or inhibitors of these
CC proteins, in order to treat BMD. The GC1C2 gene offers a simpler and
CC cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
CC age-related macular dystrophy

SQ Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 U; 0 Other;

Query Match 100.0%; Score 2429; DB 2; Length 2429;

Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;

Matches 2429; Conservative 0; Pident 0; Indels 0; Gaps 0;

QY 1 CAGGAGTCCACAGCCTAGTCGACGACCTTCTGAGGATCATGGAACCCACCTGGA 60
Db 1 CAGGAGTCCACAGCCTAGTCGACGACCTTCTGAGGATCATGGAACCCACCTGGA 60
QY 61 CCCCACTGACCCAGGCCACCTGTCGACGCCACCTGCGCATGACATCACTTACA 120
Db 61 CCCCACTGACCCAGGCCACCTGTCGACGCCACCTGCGCATGACATCACTTACA 120
QY 121 CAAGCCAGTGGCTTAATGCCCCGTTAGGCTCTTCCCGGCTGCTGCTGTCGCGG 180
Db 121 CAAGCCAGTGGCTTAATGCCCCGTTAGGCTCTTCCCGGCTGCTGCTGTCGCGG 180
QY 181 GCAGCATCTACAGCTGCTATATGCGAGTTCTTATCTTCTGCTGCTGCTACTACATCA 240
Db 181 GCAGCATCTACAGCTGCTATATGCGAGTTCTTATCTTCTGCTGCTGCTACTACATCA 240
QY 241 TCCGCTTTATTTATAGGCTGGCCCTCACGGAAACAAACAGCTGATGTTGAAATGGA 300
Db 241 TCCGCTTTATTTATAGGCTGGCCCTCACGGAAACAAACAGCTGATGTTGAAATGGA 300
QY 301 CTCTGATTGCGACACAGTCAATCAAGCTCATCCCATTTCTTCCGTCGGCTTTCTAG 360
Db 301 CTCTGATTGCGACACAGTCAATCAAGCTCATCCCATTTCTTCCGTCGGCTTTCTAG 360
QY 361 TGAAGCTGTCGTCGACCCGCTGTCGGAACCAATGAGAACTGTCGTCGACCGCC 420
Db 361 TGAAGCTGTCGTCGACCCGCTGTCGGAACCAATGAGAACTGTCGTCGACCGCC 420
QY 421 TCATGAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 480
Db 421 TCATGAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 480
QY 481 GCAAGCTCATCCGCTACCGCAACCTGCGCAACGTCCTCATCTGCGAGCGTCAGCACCG 540
Db 481 GCAAGCTCATCCGCTACCGCAACCTGCGCAACGTCCTCATCTGCGAGCGTCAGCACCG 540
QY 541 CAGTCTACAGAGCTTCCCGACGCGCCGACGACCTGTCGACAGCGCTTATGACTCCG 600
Db 541 CAGTCTACAGAGCTTCCCGACGCGCCGACGACCTGTCGACAGCGCTTATGACTCCG 600
QY 601 CAGAAACACAGCAGTGGAGAACTGAGCTTACCAACAATGTTCTGAGTGGCTGCG 660
Db 601 CAGAAACACAGCAGTGGAGAACTGAGCTTACCAACAATGTTCTGAGTGGCTGCG 660
QY 661 TGTGTTTGGCAACCTGCTCAATGAAAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTC 720
Db 661 TGTGTTTGGCAACCTGCTCAATGAAAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTC 720
QY 721 TGTCTCAGAGCTGCTGACAGAGTGAACCTTGCCTACCTGTCGTCGTCGTCGTCGTC 780
Db 721 TGTCTCAGAGCTGCTGACAGAGTGAACCTTGCCTACCTGTCGTCGTCGTCGTCGTC 780
QY 781 CTTACAGCTGATTAATGATCCCACTGCTGATTAACAAGTGTGACCTGTCGTCGTCGTC 840
Db 781 CTTACAGCTGATTAATGATCCCACTGCTGATTAACAAGTGTGACCTGTCGTCGTCGTC 840
QY 841 GCTTCTCTGACCTGCTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 900
Db 841 GCTTCTCTGACCTGCTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 900

QY 901 GCCATGACGTGACCTCTGTCGCCCCGCTTCAAGTTCCTGACGTTCTTCTATGTTG 960
Db 901 GCCATGACGTGACCTCTGTCGCCCCGCTTCAAGTTCCTGACGTTCTTCTATGTTG 960
QY 961 GCTGCTGAAAGTGGGCTCTTCCAGGGCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTC 1020
Db 961 GCTGCTGAAAGTGGGCTCTTCCAGGGCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTC 1020
QY 1021 ATGCGACAGCTGCTTGAACGAGATGATGATGATGATGATGATGATGATGATGATGATG 1080
Db 1021 ATGCGACAGCTGCTTGAACGAGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 AAAGCAGCAGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1140
Db 1081 AAAGCAGCAGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1140
QY 1141 GAGATGCTTGAACCCGCGGAGGCGGTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1200
Db 1141 GAGATGCTTGAACCCGCGGAGGCGGTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1200
QY 1201 TGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 TGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
QY 1261 CTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1320
Db 1261 CTTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1320
QY 1321 GAATTAAGCCGAGCACAAGCCCTTACAAGCTGCTTCCGCGCTGTCGTCGTCGTCGTCGTC 1380
Db 1321 GAATTAAGCCGAGCACAAGCCCTTACAAGCTGCTTCCGCGCTGTCGTCGTCGTCGTCGTC 1380
QY 1381 CTTTATGCGCTCACTTCAACATCAAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1440
Db 1381 CTTTATGCGCTCACTTCAACATCAAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1440
QY 1441 TCAGAGAGCAGAGAGATGCTCAAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1500
Db 1441 TCAGAGAGCAGAGAGATGCTCAAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1500
QY 1501 CCAATGATCAACATCTTCCAGGGCAAACTCAAGAGCAAACTGTCGTCGTCGTCGTCGTC 1560
Db 1501 CCAATGATCAACATCTTCCAGGGCAAACTCAAGAGCAAACTGTCGTCGTCGTCGTCGTC 1560
QY 1561 ATCCCTTCCAGAGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1620
Db 1561 ATCCCTTCCAGAGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1620
QY 1621 CCAAGAGAGCAAGAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1680
Db 1621 CCAAGAGAGCAAGAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1680
QY 1681 GTATCAGAGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1740
Db 1681 GTATCAGAGGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1740
QY 1741 CTTCCCTTGAACCATGAGCGCTTCAAGCTTCAAGTGTCAAGGATGACCAAA 1800
Db 1741 CTTCCCTTGAACCATGAGCGCTTCAAGCTTCAAGTGTCAAGGATGACCAAA 1800
QY 1801 AGACAAAAGCTTAAAGACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1860
Db 1801 AGACAAAAGCTTAAAGACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1860
QY 1861 GAGCGATGGGCTTGAAGTGAAGCAACCAAGATGATGATGATGATGATGATGATGATGATG 1920
Db 1861 GAGCGATGGGCTTGAAGTGAAGCAACCAAGATGATGATGATGATGATGATGATGATGATG 1920
QY 1921 GTTTAACTGACGATATGACAGATCCCGAAATCACTTCAAGAGACCTTTGAGACA 1980
Db 1921 GTTTAACTGACGATATGACAGATCCCGAAATCACTTCAAGAGACCTTTGAGACA 1980

Qy 1981 ATCAACAACAACATACACTACATCAAAAGATCATGATCCTTAT796GCGCTTGA 2040
Db 1981 ATCAACAACAACATACACTACATCAAAAGATCATGATCCTTAT796GCGCTTGA 2040
Qy 2041 AAACAGGATGAGACATTCCTTAACCTGCTTCTTAATGGGGATCTTCCGACGCAAGT 2100
Db 2041 AAACAGGATGAGACATTCCTTAACCTGCTTCTTAATGGGGATCTTCCGACGCAAGT 2100
Qy 2101 CCTGACCTGTGTGATGACAGAGAGACATGATCCAGTCACAGCCATCAGCTGTCCACA 2160
Db 2101 CCTGACCTGTGTGATGACAGAGAGACATGATCCAGTCACAGCCATCAGCTGTCCACA 2160
Qy 2161 CTGAAGAACGTCCTTCAACAGCCTGATCAATGTTAGTTAATGATTAATTAATCC 2220
Db 2161 CTGAAGAACGTCCTTCAACAGCCTGATCAATGTTAGTTAATGATTAATTAATCC 2220
Qy 2221 AGACTACTTCAAGCCTTATATGCTTTTATTCATTAATACTGTGAAGCTAGACTGAACA 2280
Db 2221 AGACTACTTCAAGCCTTATATGCTTTTATTCATTAATACTGTGAAGCTAGACTGAACA 2280
Qy 2281 TTGGAAACATTTAATCTGAGACTGAGATTCAGAGTCGGGAAACCTTATGTTCTATCTGAT 2340
Db 2281 TTGGAAACATTTAATCTGAGACTGAGATTCAGAGTCGGGAAACCTTATGTTCTATCTGAT 2340
Qy 2341 CCAAGACAGCAGCAGCCTTATGATTAATGCTCCCAAACTAATGATTAATTAATCAAAATCT 2400
Db 2341 CCAAGACAGCAGCAGCCTTATGATTAATGCTCCCAAACTAATGATTAATTAATCAAAATCT 2400
Qy 2401 CGTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2429
Db 2401 CGTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2429

RESULT 2

ADQ84435
ID ADQ84435 standard; cDNA; 2404 BP.

AC ADQ84435;

DT 07-OCT-2004 (first entry)

XX Human tumour-associated antigenic target (TAT) cDNA sequence #1249.

XX human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.

OS Homo sapiens.

XX MO2004060270-A2.

PN 22-JUL-2004.

XX 15-OCT-2003; 2003WO-US029126.

PR 18-OCT-2002; 2002US-0418988P.

XX (GETH) GENENTECH INC.

PA (WOTD/) WU T D.

PA (ZHOU/) ZHOU Y.

XX Mu TD, Zhou Y;

XX WPI; 2004-534300/51.

XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.

XX Claim 1; SEQ ID NO 1249; 5504bp; English.

XX The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%

CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a) -
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) method of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

XX SQ Sequence 2404 BP; 590 A; 696 C; 611 G; 507 T; 0 U; 0 Other;

XX Query Match 98.8%; Score 2399.2; DB 12; Length 2404;

XX Best Local Similarity 99.9%; Pred. No. 0;

XX Matches 2401; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAGGAGTCCACCAAGCTAGTGCAGACCTTCTGTGGATCATGGAACCCACTGGAA 60
Db 1 CAGGAGTCCACCAAGCTAGTGCAGACCTTCTGTGGATCATGGAACCCACTGGAA 60
Qy 61 CCCCACTGACCCAAAGCCCACTGCTGACGCCCACTGCTGGCCATGACCATTAACA 120
Db 61 CCCCACTGACCCAAAGCCCACTGCTGACGCCCACTGCTGGCCATGACCATTAACA 120
Qy 121 CAAGCAAGTGGCTAATGCGGCTTAAGCTCTTCTCCGCGCTGCTGTGAGCGGG 180
Db 121 CAAGCAAGTGGCTAATGCGGCTTAAGCTCTTCTCCGCGCTGCTGTGAGCGGG 180
Qy 181 GCAGATCTACAGCTGATATATGCGAGTTCCTTAATCTTCTGCTGCTACTACATCA 240
Db 181 GCAGATCTACAGCTGATATATGCGAGTTCCTTAATCTTCTGCTGCTACTACATCA 240
Qy 241 TCCGCTTATTTAATGAGTGGCCCTCAAGGAAACAAACACTGATGTTTGAAGAACTGA 300
Db 241 TCCGCTTATTTAATGAGTGGCCCTCAAGGAAACAAACACTGATGTTTGAAGAACTGA 300
Qy 301 CTCTGATTTGAGACAGCTACATCCAGCTCATCCCATTTCTTCTGTGCTGGGCTTCTACG 360
Db 301 CTCTGATTTGAGACAGCTACATCCAGCTCATCCCATTTCTTCTGTGCTGGGCTTCTACG 360
Qy 361 TGACGCTGTGTGTAACCGCTGTGTGAACAGTAAGAAACCTGCGGCGCCAGCGCC 420
Db 361 TGACGCTGTGTGTAACCGCTGTGTGAACAGTAAGAAACCTGCGGCGCCAGCGCC 420
Qy 421 TCATGAGCTGTGTGCTGGGCTTCTGTAAGGCAAGACAGCAAGCGGCTGCTGCGGC 480
Db 421 TCATGAGCTGTGTGCTGGGCTTCTGTAAGGCAAGACAGCAAGCGGCTGCTGCGGC 480

QY 481 GCAGGCTCATCCGCTACCGCAACCTGAGCAACGTGCTCATCTGAGGCTGACGACCG 540
Db 481 GCAGGCTCATCCGCTACCGCAACCTGAGCAACGTGCTCATCTGAGGCTGACGACCG 540
QY 541 CAGCTCTACAAAGCGCTTCCCGACGCGCCAGACCTGTGTGCAAGCAGGCTTTATGACCTCCG 600
Db 541 CAGCTCTACAAAGCGCTTCCCGACGCGCCAGACCTGTGTGCAAGCAGGCTTTATGACCTCCG 600
QY 601 CAGAACACAAAGCATTTGAGAACTGAGGCTTACCAACATGTTTGGGTGCTCCCTGAG 660
Db 601 CAGAACACAAAGCATTTGAGAACTGAGGCTTACCAACATGTTTGGGTGCTCCCTGAG 660
QY 661 TGTGTTTGGCAACCTGTCAATGAAAGCGTGTGAGGTGCAATCCGAGACCTATCC 720
Db 661 TGTGTTTGGCAACCTGTCAATGAAAGCGTGTGAGGTGCAATCCGAGACCTATCC 720
QY 721 TGCTCCAGAGCTGTGTAAGAGATGAACTGTTGGGTATCTCATGTGTGACACCTGTATG 780
Db 721 TGCTCCAGAGCTGTGTAAGAGATGAACTGTTGGGTATCTCATGTGTGACACCTGTATG 780
QY 781 CTTACGACTGGAATTAATATCCACTGTGTATACACAGGTGTGACTGTGGGGGTGTACA 840
Db 781 CTTACGACTGGAATTAATATCCACTGTGTATACACAGGTGTGACTGTGGGGGTGTACA 840
QY 841 GCTTCTTCCCTGACTTGTCTATGTTGGGGGCGAGTTTCTGAACCCAGCCAAAGGCTTACCCTG 900
Db 841 GCTTCTTCCCTGACTTGTCTATGTTGGGGGCGAGTTTCTGAACCCAGCCAAAGGCTTACCCTG 900
QY 901 GCCATGAGCTGAGACTCGTTGTGCGCGTCTCAACGTTCTGTGAGATTTCTTCTATGTTG 960
Db 901 GCCATGAGCTGAGACTCGTTGTGCGCGTCTCAACGTTCTGTGAGATTTCTTCTATGTTG 960
QY 961 GCTGCTGAAAGTGTGGGCTCTTCCAGGGCCCTGTGAGGCTGAGGAGTGGCCAGAGGGTTC 1020
Db 961 GCTGCTGAAAGTGTGGGCTCTTCCAGGGCCCTGTGAGGCTGAGGAGTGGCCAGAGGGTTC 1020
QY 1021 ATGAGCAGAGCTGTGTAAGAGAGATGCAAGTGTCAAGAAAGAAAGTGTCTACGAGTAG 1080
Db 1021 ATGAGCAGAGCTGTGTAAGAGAGATGCAAGTGTCAAGAAAGAAAGTGTCTACGAGTAG 1080
QY 1081 AAAGCAGCAGGCGTGTGAGCGCACACCTGTATCCAGCTACTGAGGAGCTGAGGAG 1140
Db 1081 AAAGCAGCAGGCGTGTGAGCGCACACCTGTATCCAGCTACTGAGGAGCTGAGGAG 1140
QY 1141 GAGATTCGCTTGAACCCGAGAGGCGGAGGTTGTGTGCAAGACGCTCAACCCCTT 1200
Db 1141 GAGATTCGCTTGAACCCGAGAGGCGGAGGTTGTGTGCAAGACGCTCAACCCCTT 1200
QY 1201 TGGAGAGATGATGATTTTGAAGACCACTGATGTGTGCAAGAAATTTGGCAGGTTC 1260
Db 1201 TGGAGAGATGATGATTTTGAAGACCACTGATGTGTGCAAGAAATTTGGCAGGTTC 1260
QY 1261 CCTGTGTGCTGTGATGATGATGACACAGACCTGCTGTGATGAGCGGACATGTACTG 1320
Db 1261 CCTGTGTGCTGTGATGATGATGACACAGACCTGCTGTGATGAGCGGACATGTACTG 1320
QY 1321 GAATTAAGCCGAGCAGACCCGCTTACAGAGTGTTCGAGCTTCCGTCGAGGCTTC 1380
Db 1321 GAATTAAGCCGAGCAGACCCGCTTACAGAGTGTTCGAGCTTCCGTCGAGGCTTC 1380
QY 1381 CTTTATGGGCTTCACTTCAACATCAGCTGTAACAAAGAGAGATGAGTTTCCAGCCCA 1440
Db 1381 CTTTATGGGCTTCACTTCAACATCAGCTGTAACAAAGAGAGATGAGTTTCCAGCCCA 1440
QY 1441 TCAGAGAGAGAGAGATGCTACGCTGGCATATGTCGCTTCTCTAGAGCTTGCAGTTC 1500
Db 1441 TCAGAGAGAGAGAGATGCTACGCTGGCATATGTCGCTTCTCTAGAGCTTGCAGTTC 1500
QY 1501 CCATGATCACCATCTCTCCAGAGGCAACTCAAGAACCAACTCTGTGAGCCCAAGAGGA 1560
Db 1501 CCATGATCACCATCTCTCCAGAGGCAACTCAAGAACCAACTCTGTGAGCCCAAGAGGA 1560
QY 1561 ATCCCTTCTCAAGAGGCGCTGCCCAAAAACAAGGAGCCCAACAGAACGTTAGGGG 1620

Db 1561 ATCCCTTCTCAAGAGGCGCTGCCCAAAAACAAGGAGCCCAACAGAACGTTAGGGG 1620
QY 1621 CCAGAGAGCAACAAGGCTGTGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGGCCACT 1680
Db 1621 CCAGAGAGCAACAAGGCTGTGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGGCCACT 1680
QY 1681 GTATCAGAGGCGCAAGCTTCAAGTGTGAGCGCTTCAAGTGTGAGCGCTTCAAGTGTG 1740
Db 1681 GTATCAGAGGCGCAAGCTTCAAGTGTGAGCGCTTCAAGTGTGAGCGCTTCAAGTGTG 1740
QY 1741 CTTCCCTTAAACATCAGCGCTTCAAGTGTGAGCGCTTCAAGTGTGAGCGCTTCAAGTGTG 1800
Db 1741 CTTCCCTTAAACATCAGCGCTTCAAGTGTGAGCGCTTCAAGTGTGAGCGCTTCAAGTGTG 1800
QY 1801 AGACAAAGCTTAAAGCTGTGAGTGTGAGGCGCAAGAAAGTTTGAATTTGCTCAGA 1860
Db 1801 AGACAAAGCTTAAAGCTGTGAGTGTGAGGCGCAAGAAAGTTTGAATTTGCTCAGA 1860
QY 1861 GAGGATGAGGCGCTTGTGATGAGGACCCAGAAATCTCAAGTGTGAGGAGAAACCTGTGA 1920
Db 1861 GAGGATGAGGCGCTTGTGATGAGGACCCAGAAATCTCAAGTGTGAGGAGAAACCTGTGA 1920
QY 1921 GTTTAACCTGACGGAATATGCAAGATCCCGGAAATCACTCAAGAAACCTTTGGAACA 1980
Db 1921 GTTTAACCTGACGGAATATGCAAGATCCCGGAAATCACTCAAGAAACCTTTGGAACA 1980
QY 1981 ATCAACAACAATACACTCACTCAAGATCATGATCTTATTTGGGCGCTTGA 2040
Db 1981 ATCAACAACAATACACTCACTCAAGATCATGATCTTATTTGGGCGCTTGA 2040
QY 2041 AAACAGGAGTGAAGCAATTTCTTAACTGCTTCTTATGAGGATGCTTGCAGCGCAGGT 2100
Db 2041 AAACAGGAGTGAAGCAATTTCTTAACTGCTTCTTATGAGGATGCTTGCAGCGCAGGT 2100
QY 2101 CTTACCTGTGTGTACACAGAGGACACTGATTCAGACGCACTGATTCAGCTGTGTACA 2160
Db 2101 CTTACCTGTGTGTACACAGAGGACACTGATTCAGACGCACTGATTCAGCTGTGTACA 2160
QY 2161 CTGAAAGAACTGTCTTCAACAAGCTGTAATCAATGTTAGTTAATGATTAATAATCCC 2220
Db 2161 CTGAAAGAACTGTCTTCAACAAGCTGTAATCAATGTTAGTTAATGATTAATAATCCC 2220
QY 2221 AGACTACTCAGCTTAAATGCTTAACTTCAATTAAGTGTGAAAGCTAGACTGAACTCA 2280
Db 2221 AGACTACTCAGCTTAAATGCTTAACTTCAATTAAGTGTGAAAGCTAGACTGAACTCA 2280
QY 2281 TTGAAAACATTTAACTCAAGCTGTGATTCAGAGTCCGGAACCTTATCTATCTGAAT 2340
Db 2281 TTGAAAACATTTAACTCAAGCTGTGATTCAGAGTCCGGAACCTTATCTATCTGAAT 2340
QY 2341 CCAAGACGCGCACACTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 2400
Db 2341 CCAAGACGCGCACACTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 2400
QY 2401 CGTT 2404
Db 2401 CGTT 2404

RESULT 3
AD083262
ID AD083262 standard; cDNA; 2404 BP.
XX AD083262;
AC
XX
DT 07-OCT-2004 (first entry)
XX
XX Human tumour-associated antigenic target (TAT) cDNA sequence #76.
DE human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KW cancer; cell proliferative disorder; gene; ss.
XX

OS Homo sapiens.
XX WO2004060270-A2.
XX 22-JUL-2004.
XX 15-OCT-2003; 2003WO-US029126.
XX 18-OCT-2002; 2002US-0418988P.
XX (GETH) GENENTECH INC.
XX (WUTD/) WU T D.
XX (ZHOU/) ZHOU Y.
XX Mu TD, Zhou Y;
XX WPI; 2004-534300/51.
XX
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
XX preventing or treating cell proliferative disorders such as cancer.
XX
XX Claim 1; SEQ ID NO 76; 5504bp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
XX target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
XX sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
XX (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
XX sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
XX (c). Also described: (1) an expression vector comprising the above
XX nucleic acid; (2) a host cell comprising the above expression vector; (3)
XX a process for producing a polypeptide; (4) an isolated polypeptide
XX comprising: (a) an amino acid sequence encoded by any of the above
XX nucleotide sequences; (b) an amino acid sequence encoded by the full-
XX length coding region of the above nucleotide sequences; or (c) a sequence
XX having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
XX comprising the above polypeptide fused to a heterologous polypeptide; (6)
XX an isolated antibody that binds to the above polypeptide; (7) a process
XX for producing the antibody; (8) an isolated oligopeptide that binds to
XX the above polypeptide; (9) a tumour-associated antigenic target (TAT)
XX binding organic molecule that binds to the above polypeptide; (10) a
XX composition of matter comprising the above (chimeric) polypeptide,
XX antibody, oligopeptide or TAT binding organic molecule, in combination
XX with a carrier; (11) an article of manufacture comprising a container and
XX the composition of matter contained within the container; (12) methods of
XX inhibiting the growth of a cell that expresses the above protein, where
XX the growth of the cell is at least in part dependent upon a growth
XX potentiating effect of the above protein; (13) a method of
XX therapeutically treating a mammal having a cancerous tumour comprising
XX cells that express the above protein; (14) a method of determining the
XX presence of a protein in a sample suspected of containing the protein
XX described above; (15) methods of diagnosing the presence of a tumour in a
XX mammal; (16) a method for treating or preventing a cell proliferative
XX disorder associated with increased expression or activity of the above
XX protein; and (17) a method of binding an antibody, oligopeptide or
XX organic molecule to a cell that expresses the protein described above.
XX The TAT sequences have cytostatic activities, and can be used in gene
XX therapy. The composition and methods are useful for diagnosing,
XX preventing or treating cancer. The composition is also used for preparing
XX a medicament for the therapeutic treatment or diagnostic detection of a
XX cell proliferative disorder or cancer. The present sequence represents a
XX human TAT cDNA sequence from the present invention.
XX
SQ Sequence 2404 BP, 590 A, 696 C, 611 G, 507 T, 0 U, 0 Other;
Query Match 98.8%; Score 2399.2; DB 13; Length 2404;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2401; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db
61 CCCACCTGACCCCAAGCCCACTGCTGAGCCCACTGCGCATGACATCACTTACA 120
Oy 121 CAAGCCAAAGTGGCTAAATGCCCCCTTAGGCTCTTCCCGGCTGCTGAGCGGG 180
Db 121 CAAGCCAAAGTGGCTAAATGCCCCCTTAGGCTCTTCCCGGCTGCTGAGCGGG 180
Oy 181 GCAGCATCTCAAGCTCTATATGCGAGTTCTTAATCTTCTGCTCTGCTACTACATCA 240
Db 181 GCAGCATCTCAAGCTCTATATGCGAGTTCTTAATCTTCTGCTCTGCTACTACATCA 240
Oy 241 TCCGCTTTATTTATAGCTGAGCCCTCAAGGAAACAACAGCTATGTTGAAAACTGA 300
Db 241 TCCGCTTTATTTATAGCTGAGCCCTCAAGGAAACAACAGCTATGTTGAAAACTGA 300
Oy 301 CTCTGTATTTGGAAGAGCTACATCCAGCTCATCCCAATTTCTTGCTGCTGGCTTACG 360
Db 301 CTCTGTATTTGGAAGAGCTACATCCAGCTCATCCCAATTTCTTGCTGCTGGCTTACG 360
Oy 361 TGAAGCTGTGTGTAAGCCGCTGCTGTAAGCAAGTACAGAACTGCTGCTGCTGCTGCTG 420
Db 361 TGAAGCTGTGTGTAAGCCGCTGCTGTAAGCAAGTACAGAACTGCTGCTGCTGCTGCTG 420
Oy 421 TCATGAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 TCATGAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Oy 481 GCAGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GCAGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Oy 541 CAGCTCTCAAGCGCTTCCCGGCGCCGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 CAGCTCTCAAGCGCTTCCCGGCGCCGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Oy 601 CAGAACCAAGCACTGTGAGAACTGAAGCTTACCAACAACATGTTCTGAGTCCCTGG 660
Db 601 CAGAACCAAGCACTGTGAGAACTGAAGCTTACCAACAACATGTTCTGAGTCCCTGG 660
Oy 661 TGTGTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 TGTGTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Oy 721 TGTGCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 TGTGCTGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Oy 781 CCTACGACTGATTAATATCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 CCTACGACTGATTAATATCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Oy 841 GCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 GCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Oy 901 GGCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 GGCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Oy 961 GCTGCTGTAAGTGGGCTCTCCAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 961 GCTGCTGTAAGTGGGCTCTCCAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Oy 1021 ATGAGCCAGAGCTCTGTAAGAGAGATGAGTGTGAGAGAAAGAGTCTTACAGGGTATG 1080
Db 1021 ATGAGCCAGAGCTCTGTAAGAGAGATGAGTGTGAGAGAAAGAGTCTTACAGGGTATG 1080
Oy 1081 AAAGCAGCAGGCGTGTGAGGAGCAGACCTGTAATCCAGCTACTCGGAGGCTGAGGAG 1140
Db 1081 AAAGCAGCAGGCGTGTGAGGAGCAGACCTGTAATCCAGCTACTCGGAGGCTGAGGAG 1140
Oy 1141 GAGATGCTTTGAAACCGGAGGCGGAGGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 1141 GAGATGCTTTGAAACCGGAGGCGGAGGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200

```

Db      1141 GAGAAATCGCTTGAAACCCGAGAGCGAGGTTGTGTGGCAGAGAGCTCATCAACCCCTT 1200
Qy      1201 TGGAGAGATGATGATGATTTTGGAGACCACTGGATTGTTCAGAGAAATTTTGCAGGTTC 1260
Db      1201 TGGAGAGATGATGATGATTTTGGAGACCACTGGATTGTTCAGAGAAATTTTGCAGGTTC 1260
Qy      1261 CCTGTGGCTGTGTGATGATGACACAGAGACCTGCTCGATGAGAGCCGAGCATGTACTG 1320
Db      1261 CCTGTGGCTGTGTGATGATGACACAGAGACCTGCTCGATGAGAGCCGAGCATGTACTG 1320
Qy      1321 GAATTAAGCCGAGGACAGGCCCCCTTACAGAGCTGCTTCGGCCAGATTCCGTGAGCTTC 1380
Db      1321 GAATTAAGCCGAGGACAGGCCCCCTTACAGAGCTGCTTCGGCCAGATTCCGTGAGCTTC 1380
Qy      1381 CTTTATGGGCTCCACCTTCAACATCACTGATGAAACAAAGAGAGATGAGATTCCAGCCCA 1440
Db      1381 CTTTATGGGCTCCACCTTCAACATCACTGATGAAACAAAGAGAGATGAGATTCCAGCCCA 1440
Qy      1441 TCAGAGAGAGAGAGAGATGCTCAGCTGGCATCTATTGCGCTTCTTAGGCTTGCAATC 1500
Db      1441 TCAGAGAGAGAGAGAGATGCTCAGCTGGCATCTATTGCGCTTCTTAGGCTTGCAATC 1500
Qy      1501 CCATGATCACCATCTCTCCAGGGGCAAACTCAAGAACAACTAATCTGTGGCCCAAGAGGA 1560
Db      1501 CCATGATCACCATCTCTCCAGGGGCAAACTCAAGAACAACTAATCTGTGGCCCAAGAGGA 1560
Qy      1561 ATCCCTTCTCAAGAGGCTCTGCCCAAAAACAAGAGCCCAACAGAAAGCTTGAAGGG 1620
Db      1561 ATCCCTTCTCAAGAGGCTCTGCCCAAAAACAAGAGCCCAACAGAAAGCTTGAAGGG 1620
Qy      1621 CCAGAGAGACAAAGGCTCTGAAAGCTTAAAGCTGTGAGAGCGCTTAAATCTTGAGCCCA 1680
Db      1621 CCAGAGAGACAAAGGCTCTGAAAGCTTAAAGCTGTGAGAGCGCTTAAATCTTGAGCCCA 1680
Qy      1681 GATATCAGAGGCGCAGGCTACTACAGTGCCTCCACAGAGCGCTTCAAGCTGAGCCCA 1740
Db      1681 GATATCAGAGGCGCAGGCTACTACAGTGCCTCCACAGAGCGCTTCAAGCTGAGCCCA 1740
Qy      1741 CTTCCCCCTGAAACCATCAAGCGCTGAAAGCTTCAAGTGTCAAGGATGAGACACCA 1800
Db      1741 CTTCCCCCTGAAACCATCAAGCGCTGAAAGCTTCAAGTGTCAAGGATGAGACACCA 1800
Qy      1801 AGACAAAGACCTTAAAGACTGTGAGTCTGAGGCGCAAGAAAGCTTGAATGCTCTCAGA 1860
Db      1801 AGACAAAGACCTTAAAGACTGTGAGTCTGAGGCGCAAGAAAGCTTGAATGCTCTCAGA 1860
Qy      1861 GAGCGATGGGCGCTTGTGATGAGACACCCAGAAAGTATCTCAAGTGAAGAGAAAGCTG 1920
Db      1861 GAGCGATGGGCGCTTGTGATGAGACACCCAGAAAGTATCTCAAGTGAAGAGAAAGCTG 1920
Qy      1921 GTTTAACTGACGAGATATGCGAGATATCCGAGAAATCACTTCAAGAAAGCTTTGGA 1980
Db      1921 GTTTAACTGACGAGATATGCGAGATATCCGAGAAATCACTTCAAGAAAGCTTTGGA 1980
Qy      1981 ATCACAACCAACATACACACTCACTCAAGATCACTGATCTTATTTGGGCTTGA 2040
Db      1981 ATCACAACCAACATACACACTCACTCAAGATCACTGATCTTATTTGGGCTTGA 2040
Qy      2041 AAACGGAGATGAAGACATTTCTTAACCTGCTTCTTAATGGGAGATGCTTCCGACAGGT 2100
Db      2041 AAACGGAGATGAAGACATTTCTTAACCTGCTTCTTAATGGGAGATGCTTCCGACAGGT 2100
Qy      2101 CCTCACTGTGTGACACAGAGAGACATGATCAGTCAAGCAAGCAATACAGTGTCCACA 2160
Db      2101 CCTCACTGTGTGACACAGAGAGACATGATCAGTCAAGCAAGCAATACAGTGTCCACA 2160
Qy      2161 CTGAAGAAAGCTGTCTTCAACAGCCTGAATCAAAATGTTAGCTTAATAGATAAAATCCC 2220
Db      2161 CTGAAGAAAGCTGTCTTCAACAGCCTGAATCAAAATGTTAGCTTAATAGATAAAATCCC 2220
Qy      2221 AGACTACTTCAAGCTTTAAATGCTTTTATTCATATAAAAGCTGAAGCTGAACCA 2280
Db      2221 AGACTACTTCAAGCTTTAAATGCTTTTATTCATATAAAAGCTGAAGCTGAACCA 2280

```

```

Qy      2281 TTGAAACATTTAACTAGACTGTGATTCAGAGTGGGAACCTTAGTTCTATCTGAT 2340
Db      2281 TTGAAACATTTAACTAGACTGTGATTCAGAGTGGGAACCTTAGTTCTATCTGAT 2340
Qy      2341 CCAAGACGCCACACTTAGTATCTGCGCAAACTAATGATTTAATTAATAAATACT 2400
Db      2341 CCAAGACGCCACACTTAGTATCTGCGCAAACTAATGATTTAATTAATAAATACT 2400
Qy      2401 CGTT 2404
Db      2401 CGTT 2404

RESULT 4
AAZ21227
ID AAZ21227 standard; cDNA; 2229 BP.
XX
XX AAZ21227;
AC
XX
XX 22-NOV-1999 (first entry)
DT
XX
XX Human CG1CE short form cDNA sequence.
DE
XX
XX CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
KW age-related macular dystrophy; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 105..1862
FT
FT CDS
FT
FT
FT
FT W09943695-A1.
XX
XX 02-SEP-1999.
PD
XX
XX 22-FEB-1999; 99MO-US003790.
PF
XX
XX 25-FEB-1998; 98US-0075941P.
PR 18-DEC-1998; 98US-0112926P.
XX
XX (MERI ) MERCK & CO INC.
PA (UYUP-) UNIV UPPSALA.
PA
XX
XX Petrukhin K, Caskey CT, Metzker M, Wadelius C;
PI
XX
XX WPI; 1999-540560/45.
DR
XX
XX P-PSDB; AAY29953.
DR
XX
XX Human and mouse polynucleotides encoding CG1CE polypeptides.
PT
XX
XX Claim 2; Fig 2; 67pp; English.
PS
XX
XX The present sequence represents the human CG1CE cDNA sequence, which when
CC mutated is responsible for Best's macular dystrophy (BMD).
CC Polynucleotides encoding CG1CE are useful for diagnosing whether a
CC patient carries a mutation in the CG1CE gene. Normal and mutated CG1CE
CC proteins are useful for identifying activators and/or inhibitors of these
CC proteins, in order to treat BMD. The CG1CE gene offers a simpler and
CC cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
CC age-related macular dystrophy
CC
XX
SQ Sequence 2229 BP; 575 A; 646 C; 532 G; 476 T; 0 U; 0 Other;

Query Match 82.9%; Score 2013; DB 2; Length 2229;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 2226; Conservative 0; Mismatches 0; Indels 203; Gaps 1;

Qy      1 CAGGAGTCCACACAGCTAGTGCAGACCTTGTGGATCATGGACCACTTGAA 60

```

Db 1 CAGGAGTCCACGAGCTAGTCGACAGACCTTGTGGGATCATCGAACCCACCTGGAA 60
Qy 61 CCCACCTGACCCAGCCCACTGCTGAGAGCCCACTGCGCCATGACATCACTTACA 120
Db 61 CCCACCTGACCCAGCCCACTGCTGAGAGCCCACTGCGCCATGACATCACTTACA 120
Qy 121 CAAGCAAGTGGCTAATGCCGCTTAGGCTCTTCTCCGCGCTGCTGCTGCTGCTGCGAG 180
Db 121 CAAGCAAGTGGCTAATGCCGCTTAGGCTCTTCTCCGCGCTGCTGCTGCTGCTGCGAG 180
Qy 181 GCAGCATCTAACAGTGTCTAATATGCGAGATTCCTAATCTTCTGCTCTGCTAATCA 240
Db 181 GCAGCATCTAACAGTGTCTAATATGCGAGATTCCTAATCTTCTGCTCTGCTAATCA 240
Qy 241 TCCGCTTATTTAATAGCTGCGCCCTCAGCGAAGAACAAAGCTGATGTTTGAAGAACTGA 300
Db 241 TCCGCTTATTTAATAGCTGCGCCCTCAGCGAAGAACAAAGCTGATGTTTGAAGAACTGA 300
Qy 301 CTCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCTAG 360
Db 301 CTCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTCTAG 360
Qy 361 TGAAGCTGGTGTGACCCGCTGTGGAAACAGTACAGAACTGCGCTGTGCGACCGCC 420
Db 361 TGAAGCTGGTGTGACCCGCTGTGGAAACAGTACAGAACTGCGCTGTGCGACCGCC 420
Qy 421 TCATGAGCTGGTGTGCGGCTTCTGCGAAGGCAAGAAAGCAAGCCGTGCTGCTGCGGC 480
Db 421 TCATGAGCTGGTGTGCGGCTTCTGCGAAGGCAAGAAAGCAAGCCGTGCTGCTGCGGC 480
Qy 481 GCAGCTCATCGCTACGCGCAACCTGCGGACAGTCTCATCTTGGCGACGTCAGACACG 540
Db 481 GCAGCTCATCGCTACGCGCAACCTGCGGACAGTCTCATCTTGGCGACGTCAGACACG 540
Qy 541 CAGTCTACAGCGCTTCCCAAGCGCCAGCACTTGTGTCAAGCAGAGCTTATGACTCCGG 600
Db 541 CAGTCTACAGCGCTTCCCAAGCGCCAGCACTTGTGTCAAGCAGAGCTTATGACTCCGG 600
Qy 601 CAGAACACAGAGTGGAGAACTGAGCTTACCAACAACATGTTCTGGGTGCGCTGGG 660
Db 601 CAGAACACAGAGTGGAGAACTGAGCTTACCAACAACATGTTCTGGGTGCGCTGGG 660
Qy 661 TGTGGTTTGGCAACCTGTCAATGAAAGGCGTGTGAGGTGCAATCCGGGACCTATCC 720
Db 661 TGTGGTTTGGCAACCTGTCAATGAAAGGCGTGTGAGGTGCAATCCGGGACCTATCC 720
Qy 721 TGTCTCAGAGCGCTGTGAAGAGATGAACACCTTGCCTACTCAGTGTGAGACACTGTATG 780
Db 721 TGTCTCAGAGCGCTGTGAAGAGATGAACACCTTGCCTACTCAGTGTGAGACACTGTATG 780
Qy 781 CCTACAGACTGATATGATATCCCATCTGTGTATACACAGGTGTGATCTGTGCGGTGTACA 840
Db 781 CCTACAGACTGATATGATATCCCATCTGTGTATACACAGGTGTGATCTGTGCGGTGTACA 840
Qy 841 GCTTCTTCTCACTGTCTAGTGTGGGCGGAGTTCTGAACCCAGCCCAAGGCTTACCTCG 900
Db 841 GCTTCTTCTCACTGTCTAGTGTGGGCGGAGTTCTGAACCCAGCCCAAGGCTTACCTCG 900
Qy 901 GCGATGAGCTGAGCTCGTGTGCGCCGCTTCAAGCTTCTGAGATTTCTTCTTCTATATGTTG 960
Db 901 GCGATGAGCTGAGCTCGTGTGCGCCGCTTCAAGCTTCTGAGATTTCTTCTTCTATATGTTG 960
Qy 961 GCTGCTGAGAGTGGGCTCTTCCAGGAGCCCTGCTGGGCTGAGGCAATGCCCCAGAGGGGTC 1020
Db 961 GCTGCTGAGAGTGGGCTCTTCCAGGAGCCCTGCTGGGCTGAGGCAATGCCCCAGAGGGGTC 1020
Qy 971 GCTGCTGAGAGTGGGCTCTTCCAGGAGCCCTGCTGGGCTGAGGCAATGCCCCAGAGGGGTC 970
Db 971 GCTGCTGAGAGTGGGCTCTTCCAGGAGCCCTGCTGGGCTGAGGCAATGCCCCAGAGGGGTC 970
Qy 1081 AATGGCAGACAGCTGCTTGAAGAGAGATGACAGTGTCAAGAAAGAAAGTCTCACGGGTAG 1080
Db 1081 AATGGCAGACAGCTGCTTGAAGAGAGATGACAGTGTCAAGAAAGAAAGTCTCACGGGTAG 1080
Qy 971 ----- 970
Db 971 ----- 970

Db 971 ----- 970
Qy 1141 GAGAAATGCTTGAACCCGGAGGCGGAGTTGTGGTGGCAGAGCAGCTCATCAACCCCTT 1200
Db 971 -----GCTGGCAGAGCAGCTCATCAACCCCTT 997
Qy 1201 TGAAGAGATGATGATGATTTTGAAGACCAACTGATTTGTGACAGAAATTTGACAGTGTCT 1260
Db 998 TGAAGAGATGATGATGATTTTGAAGACCAACTGATTTGTGACAGAAATTTGACAGTGTCT 1057
Qy 1261 CCTGTTGGCTGTGATGATGATGACACAGAGACTGCTGTGATGAGAGCCGAGATGACTG 1320
Db 1058 CCTGTTGGCTGTGATGATGATGACACAGAGACTGCTGTGATGAGAGCCGAGATGACTG 1117
Qy 1321 GAATPAGCCGAGCCAGCCGCCCTTACACAGCTGCTTCCGCCGAGTTCCGTCGAGCCTC 1380
Db 1118 GAATPAGCCGAGCCAGCCGCCCTTACACAGCTGCTTCCGCCGAGTTCCGTCGAGCCTC 1177
Qy 1381 CTTTATGAGCTTCAACCTTCAACATCAGCTGTAACAAAGAGAGATGAGATTCCAGCCCA 1440
Db 1178 CTTTATGAGCTTCAACCTTCAACATCAGCTGTAACAAAGAGAGATGAGATTCCAGCCCA 1237
Qy 1441 TCAGAGAGACGAGAGAGATGCTCACAGCTGCGATCATTTGGCGCTTCTTATGCGCTGCA 1500
Db 1238 TCAGAGAGACGAGAGAGATGCTCACAGCTGCGATCATTTGGCGCTTCTTATGCGCTGCA 1297
Qy 1501 CCATGATACCATCTCCCAAGGCAAACTCAAGAGCAAACTAAGTGGCCCAAGAGGGA 1560
Db 1298 CCATGATACCATCTCCCAAGGCAAACTCAAGAGCAAACTAAGTGGCCCAAGAGGGA 1357
Qy 1561 ATCCCTTCTCCAGAGGCGCTGCCCCAAAACCAAGAGCAGCAAAAGTATGAGGGG 1620
Db 1358 ATCCCTTCTCCAGAGGCGCTGCCCCAAAACCAAGAGCAGCAAAAGTATGAGGGG 1417
Qy 1621 CCAGAGAGACAAAGAGGCTGGAAGCTTAAAGGCTGTGAGCCCTTCAAGTCTGCGCCACT 1477
Db 1418 CCAGAGAGACAAAGAGGCTGGAAGCTTAAAGGCTGTGAGCCCTTCAAGTCTGCGCCACT 1477
Qy 1681 GTATCAAGAGGCGAGGCTATCAATGCGCCCAAGAGGCGCCCTGAGCCCACTCCCATGTT 1740
Db 1478 GTATCAAGAGGCGAGGCTATCAATGCGCCCAAGAGGCGCCCTGAGCCCACTCCCATGTT 1537
Qy 1741 CTTCCCCCTTAAGAACATCAGGCGCGTCAAAAGCTTCAAGTGTCAAGGCAATGAGACCAA 1800
Db 1538 CTTCCCCCTTAAGAACATCAGGCGCGTCAAAAGCTTCAAGTGTCAAGGCAATGAGACCAA 1597
Qy 1801 AGACAAAAGCTTAAAGACTGTGAGTTGTGGGCGCAAGAAAAGTTTGAATTGCTCTCAGA 1860
Db 1598 AGACAAAAGCTTAAAGACTGTGAGTTGTGGGCGCAAGAAAAGTTTGAATTGCTCTCAGA 1657
Qy 1861 GAGGATGAGGCGCTTGTATGAGACCCAGAAAGTATCTCAATGAGAGAGAAAAGTGTGGA 1920
Db 1658 GAGGATGAGGCGCTTGTATGAGACCCAGAAAGTATCTCAATGAGAGAGAAAAGTGTGGA 1717
Qy 1921 GTTTAACCTGACGGATATGACAGAGATCCCGGAAATCACTCAAAAGAACCTTTGGAACA 1980
Db 1718 GTTTAACCTGACGGATATGACAGAGATCCCGGAAATCACTCAAAAGAACCTTTGGAACA 1777
Qy 1981 ATCAACCAACCAACATACACTACATCACTCAAAAGATCACTGAGATCTTATTTGGGCTTGA 2040
Db 1778 ATCAACCAACCAACATACACTACATCACTCAAAAGATCACTGAGATCTTATTTGGGCTTGA 1837
Qy 2041 AAACAGGATGATACCACTTCTTAACTGTCTTCTTAAATGAGGATGCTTCCGACAGCAGT 2100
Db 1838 AAACAGGATGATACCACTTCTTAACTGTCTTCTTAAATGAGGATGCTTCCGACAGCAGT 1897
Qy 2101 CCTCACCTGTGTGACACAGCAGAGACACTGATCAAGTCAAGCAGCATCAAGCTGTCCACA 2160
Db 1898 CCTCACCTGTGTGTACACAGCAGAGACACTGATCAAGTCAAGCAGCATCAAGCTGTCCACA 1957
Qy 2161 CTGAAGAAAGTGTCTTCAACAGCCTGATCAATGATGATGATGATGATGATGATGATGATGAT 2220
Db 1958 CTGAAGAAAGTGTCTTCAACAGCCTGATCAATGATGATGATGATGATGATGATGATGATGAT 2017

QY	1065	AAGGCTCAGCGGGTAGAAGACGACGAGCGTGTGTGGCCACACCTGTAACTCCAGCTACT	1124
Db	867	-----	866
QY	1125	CGGAGCGCTGAGGCGAGGAATCGCTTGAAACCCGGAGGCGGAGGTTGTGTGCGAGAC	1184
Db	867	-----GGTGGCAGACG	877
QY	1185	AGCTCATCAACCCCTTTGGAGAGCATGTGATGATTTTGAACCACTGATTTGTCCACA	1244
Db	878	AGCTCATCAACCCCTTTGGAGAGCATGTGATGATTTTGAACCACTGATTTGTCCACA	937
QY	1245	GGAAATTTGCSAGGTGTCCCTGTGTGGCTGTGATGATGATGACACAGGACCTGCTCGATATG	1304
Db	938	GGAAATTTGCSAGGTGTCCCTGTGTGGCTGTGATGATGATGACACAGGACCTGCTCGATATG	997
QY	1305	AGCCGGAATGATCTGTGAATTAAGCCCGAGCCACAGCCCCCTTACACAGCTGTCTCCGCC	1364
Db	998	AGCCGGAATGATCTGTGAATTAAGCCCGAGCCACAGCCCCCTTACACAGCTGTCTCCGCC	1057
QY	1365	AGTTCCGTCGAGCCTCTTTATGTGGCTTCACCTTCAATCAAGCTTGAACAAAGAGAGA	1424
Db	1058	AGTTCCGTCGAGCCTCTTTATGTGGCTTCACCTTCAATCAAGCTTGAACAAAGAGAGA	1117
QY	1425	TGAGATTTCAGACCCCAATCAGGAGGACGAGAGGATGCTCAGCTGTGGACTCAATTGGCCGCT	1484
Db	1118	TGAGATTTCAGACCCCAATCAGGAGGACGAGAGGATGCTCAGCTGTGGACTCAATTGGCCGCT	1177
QY	1485	TCCTAGGCTTCAGTCCCATGATATCAACATCTCTCCAGGGCAAACTCAAGACCAAACTAC	1544
Db	1178	TCCTAGGCTTCAGTCCCATGATATCAACATCTCTCCAGGGCAAACTCAAGACCAAACTAC	1237
QY	1545	TGTGGCCCAAGAGGAATCCCTTCTCCACAGAGGCTGTGCCAAAAACAACAAGGACGCCA	1604
Db	1238	TGTGGCCCAAGAGGAATCCCTTCTCCACAGAGGCTGTGCCAAAAACAACAAGGACGCCA	1297
QY	1605	AACAGAACTGTAGGGGCGAGGAAGACAACAAGGCTGTGAAGCTTAAAGCTGTGACGCGCT	1664
Db	1298	AACAGAACTGTAGGGGCGAGGAAGACAACAAGGCTGTGAAGCTTAAAGCTGTGACGCGCT	1357
QY	1665	TCAGTCTGTGCCCCACTGTATCAGAGGCGCAGGCTACTACAGTGTGCCCAAGCGCCCTCA	1724
Db	1358	TCAAGTCTGTGCCCCACTGTATCAGAGGCGCAGGCTACTACAGTGTGCCCAAGCGCCCTCA	1417
QY	1725	GCCCCACTGCCATGTCTTCCCTCCCTAGAACCATCAGCGCCGTCMAAGCTTCAAGTGTCA	1784
Db	1418	GCCCCACTGCCATGTCTTCCCTCCCTAGAACCATCAGCGCCGTCMAAGCTTCAAGTGTCA	1477
QY	1785	CAGGCAATGACACCAAAAGACAAGAAAGCTTAAAGACTGTAGTTCTGGGGCCAAAGAAAGTT	1844
Db	1478	CAGGCAATGACACCAAAAGACAAGAAAGCTTAAAGACTGTAGTTCTGGGGCCAAAGAAAGTT	1537
QY	1845	TTGAATGTGCTCTCAGAGAGCGATGGGGCTTTGATGTGAGACCCAGAGATATCTCAAGTGA	1904
Db	1538	TTGAATGTGCTCTCAGAGAGCGATGGGGCTTTGATGTGAGACCCAGAGATATCTCAAGTGA	1597
QY	1905	GGAGAGAAACGTGTGAGTTTAACCTGACGGATATGGCCAGAGATCCCGCAAAATCACCTCA	1964
Db	1598	GGAGAGAAACGTGTGAGTTTAACCTGACGGATATGGCCAGAGATCCCGCAAAATCACCTCA	1657
QY	1965	AAGAACTTTTGGAAACAATCACCAACCAACATACACTACACTCACTCAAAAGTACATGTGATC	2024
Db	1658	AAGAACTTTTGGAAACAATCACCAACCAACATACACTACACTCACTCAAAAGTACATGTGATC	1717
QY	2025	CTTATTTGGGCTTTGGAAAAACGGGATGAAGCACTTTCTTAA	2065
Db	1718	CTTATTTGGGCTTTGGAAAAACGGGATGAAGCACTTTCTTAA	1758

RESULT 6
AAV99722
ID AAV99722 standard; cDNA; 1263 BP.

XX	AAV99722;
AC	
XX	
DT	26-APR-1999 (first entry)
XX	
XX	
DE	Human adult retina secreted protein bk112_15 cDNA.
XX	
KW	Secreted protein; human; retina; bk112_15; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	100..885
CDS	/*tag= a
XX	
PN	W09856909-A2.
XX	
PD	17-DEC-1998.
XX	
PF	08-JUN-1998; 98WO-US011822.
XX	
PR	11-JUN-1997; 97US-00873218.
PR	05-JUN-1998; 98US-00092722.
XX	
XX	(GEMV) GENETICS INST INC.
PA	
XX	
PI	Jacobs K, Mccoy JM, Lavallic ER, Racie LA, Treacy M, Spaulding V;
PI	Agostino MJ, Howes SH, Fachtel K;
XX	
DR	WPI: 1999-080899/07.
XX	P-PSDB; AAW95345.
PT	
PT	New polynucleotides encoding secreted human proteins - derived from human
PT	foetal brain, adult testes, foetal kidney, adult thyroid or adult retina
PT	cDNA libraries.
XX	
PS	Claim 12; Page 71-72; 113pp; English.
XX	
XX	This is the nucleotide sequence of cDNA clone bk112_15, which includes an
CC	open reading frame for a 261-amino acid polypeptide (see AAW95345). The
CC	clone was isolated from a human adult retina cDNA library using methods
CC	which are selective for cDNAs encoding secreted proteins, or was
CC	identified as encoding a secreted or transmembrane protein on the basis
CC	of computer analysis of the amino acid sequence of the encoding protein.
CC	Database searches indicate some sequence similarity to known sequences.
CC	The invention provides cDNA clones (see AAV9721-33) from human adult
CC	thyroid, adult retina, adult testis, foetal kidney and foetal brain that
CC	encode novel secreted proteins (see AAW9534-53). Each clone is
CC	individually available from deposit clone ATCC 98451 (see also AAV9734-
CC	43). The isolated polynucleotides (PNS) and proteins are predicted to
CC	have activities which would make them suitable for treating, preventing
CC	or ameliorating medical conditions in humans and animals, although no
CC	supporting data is given. Suggested activities include nutritional,
CC	cytokine, cell proliferation or differentiation, immune stimulating (e.g.
CC	as vaccines) or immune suppressing, haematopoiesis regulating, tissue
CC	growth, activin/inhibin, chemotactic/chemokinetic, haemostatic,
CC	thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour invasion
CC	suppressor, and tumour inhibition activities. The PNS are also stated to
CC	be useful for gene therapy
XX	
Sequence	1263 BP; 397 A; 342 C; 272 G; 252 T; 0 U; 0 Other;

	Query Match	Score	1244.2	DB 2	Length	1263
Match Local Similarity	99.88	Prod. No.	46-274			
Resets 1246	Conservative	0	Mismatches	5	Indels	0
Gaps						
QY	1181	GAGCAGCTCATCAACCCCTTTGGAGAGAGATGATGATTTTGGAGACCACTGGATTGTC	1240			
Db	1	GAGAGCTCATCAACCCCTTTGGAGAGATGATATATATTTTGGAGCCCACTGGATTGTC	60			
QY	1241	GACAGGATTTGGCAGGTCCTCCCTTTGGCTGTGATGATGACACAGGACTCTGCTCGG	1300			
Db	61	GACAGGATTTGGCAGGTCCTCCCTTTGGCTGTGATGATGACACAGGACTCTGCTCGG	120			

```

QY 1301 ATGAGCCGACATGTACTGGAATAGCCGAGCCGAGCCGCTACAGAGCTTCTCC 1360
DB 121 ATGAGCCGAGCATATCTACTGGAATAGCCGAGCCGAGCCGCTACAGAGCTTCTCC 180
QY 1361 GCCGATTCCTCGAGCTCTCTTTATGAGGCTCACTTCAACATCACCTGAACAAAGAG 1420
DB 181 GCCGATTCCTCGAGCTCTCTTTATGAGGCTCACTTCAACATCACCTGAACAAAGAG 240
QY 1421 GAGATGAGAGTTCAGGCCCAATCAGAGAGAGAGAGATGTCTACGCTGGCATCATTTGGC 1480
DB 241 GAGATGAGAGTTCAGGCCCAATCAGAGAGAGAGAGATGTCTACGCTGGCATCATTTGGC 300
QY 1481 CGCTTCTAGGCTCGAGCTCCATGATCACCATCTCCAGGGCAAACTCAAGAGACCAA 1540
DB 301 CGCTTCTAGGCTCGAGCTCCATGATCACCATCTCCAGGGCAAACTCAAGAGACCAA 360
QY 1541 CTACTGTGGCCCAAGAGAGAGATCCCTTCTCAAGAGGCTGCCCCAAAAACAAGAGCA 1600
DB 361 CTACTGTGGCCCAAGAGAGAGATCCCTTCTCAAGAGGCTGCCCCAAAAACAAGAGCA 420
QY 1601 GCCAAACAGAACGTTAGGGGCGCAGAAACAACAAGGCTGAAGGTTAAGGCTGTGAC 1660
DB 421 GCCAAACAGAACGTTAGGGGCGCAGAAACAACAAGGCTGAAGGCTGTGTGAC 480
QY 1661 GCCTTCAAGTCTGGCCCACTGTATCAAGAGCCAGGCTACTTACAGTCCCCAGAGCGCC 1720
DB 481 GCCTTCAAGTCTGGCCCACTGTATCAAGAGCCAGGCTACTTACAGTCCCCAGAGCGCC 540
QY 1721 CTGAGCCCACTCCCATGTTCTTCCCCCTAGAACCATCAAGCCCGTCAAGGTTTCAAGT 1780
DB 541 CTGAGCCCACTCCCATGTTCTTCCCCCTAGAACCATCAAGCCCGTCAAGGTTTCAAGT 600
QY 1781 GTACACGAGCATACACCAACAACAACCTTAAGAGCTGTAGTTCTGGGCGCAAGAA 1840
DB 601 GTACACGAGCATACACCAACAACAACCTTAAGAGCTGTAGTTCTGGGCGCAAGAA 660
QY 1841 AGTTTGAATGTCTCTCAGAGAGGAGTGGGGCTTGTAGAGCACTCCGAAAGTATCTCAA 1900
DB 661 AGTTTGAATGTCTCTCAGAGAGGAGTGGGGCTTGTAGAGCACTCCGAAAGTATCTCAA 720
QY 1901 GTGAGAGAGAAAACTGTGTGAGTTTAACTGACGATATGCCAGAGATCCCCGAAATAC 1960
DB 721 GTGAGAGAGAAAACTGTGTGAGTTTAACTGACGATATGCCAGAGATCCCCGAAATAC 780
QY 1961 CTCAAAGAACCTTTGGAAACAATCACCAACAACATACACTCACTCAAAAGATCACATG 2020
DB 781 CTCAAAGAACCTTTGGAAACAATCACCAACAACATACACTCACTCAAAAGATCACATG 840
QY 2021 GATCCTTATTTGGGCTTTGGAAAAAGAGATGAGACATTTCTTAACCTGCTTCTATGG 2080
DB 841 GATCCTTATTTGGGCTTTGGAAAAAGAGATGAGACATTTCTTAACCTGCTTCTATGG 900
QY 2081 GGAATGCTTGGCAGCGAGAGTCTCTCACTGTGTATACACAGACAGAGACATGATCAGTCA 2140
DB 901 GGAATGCTTGGCAGCGAGAGTCTCTCACTGTGTATACACAGACAGAGACATGATCAGTCA 960
QY 2141 CAGGCAATACAGCTGTCCACATGAAAGACGTTCTCAACAGCTGAAATCAATGCTTA 2200
DB 961 CAGGCAATACAGCTGTCCACATGAAAGACGTTCTCAACAGCTGAAATCAATGCTTA 1020
QY 2201 GCTTAAATGATAAAAATCCAGACTTCAAGCTTTTAAATGCTTTTATTCATTAAGAACT 2260
DB 1021 GCTTAAATGATAAAAATCCAGACTTCAAGCTTTTAAATGCTTTTATTCATTAAGAACT 1080
QY 2261 GTGAAAGCTAGACTGAGCACTTGGAAACATTTAATCAGCTCTGATTCAGAGTGGGA 2320
DB 1081 GTGAAAGCTAGACTGAGCACTTGGAAACATTTAATCAGCTCTGATTCAGAGTGGGA 1140
QY 2321 ACCCTTAAGTTCTATCTGAATCCAAGACAGCACTTATGATATCGCCCAACTAATGA 2380
DB 1141 ACCCTTAAGTTCTATCTGAATCCAAGACAGCACTTATGATATCGCCCAACTAATGA 1200

```

```

QY 2381 GTTATTAATTAATCAATACTCGTTAAAAAATTTTTTTTTTTTTTTTTTTTTT 2429
DB 1201 GTTATTAATTAATCAATACTCGTTAAAAAATTTTTTTTTTTTTTTTTTTTTT 1249

RESULT 7
ADA44960
ID ADA44960 standard; cDNA, 1238 BP.
XX
AC ADA44960;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human polynucleotide #2.
XX
KW Human; gene; ss; genetic disorder; genetic fingerprinting;
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW insulin dependent diabetes mellitus; graft-versus-host disease; anaemia;
KW periodontal disease; bone fracture; cartilage damage;
KW central nervous system disorder; Alzheimer's disease;
KW Parkinson's disease; cancer; nutrition; carbon source; nitrogen source;
KW carbohydrate source.
XX
OS Homo sapiens.
XX
PN US2003044935-A1.
XX
PD 06-MAR-2003.
XX
PE 21-DEC-2000; 2000US-00746783.
XX
PR 11-JUN-1997; 97US-0086236P.
PR 12-JUN-1997; 97US-0086234P.
PR 08-JUL-1997; 97US-0092115P.
PR 08-SEP-1997; 97US-0093045P.
PR 02-OCT-1997; 97US-0090100P.
PR 27-OCT-1997; 97US-00958304.
PR 07-NOV-1997; 97US-0090111P.
PR 05-JUN-1998; 98US-0009722.
PR 11-JUN-1998; 98US-00096287.
PR 11-JUN-1998; 98US-00098588.
PR 04-AUG-1998; 98US-00130189.
PR 08-SEP-1998; 98US-00149633.
PR 01-OCT-1998; 98US-00165960.
PR 04-NOV-1998; 98US-00185936.
XX
PA (JACO/) JACOBS K.
PA (MCCO/) MCCOY J M.
PA (LVAL/) LA VALLIE E R.
PA (COLL/) COLLINS-RACIE L A.
PA (EVAN/) EVANS C.
PA (MERB/) MERBERG D.
PA (TREA/) TREACY M.
PA (SPAU/) SPAULDING V.
XX
PI Jacobs K, McCoy JM, La Vallie ER, Collins-Racie LA, Evans C;
PI Werberg D, Treacy M, Spaulding V;
PI MPI; 2003-521754/49.
XX
DR P-PSDB; ADA44961.
XX
PT New polypeptides and polynucleotides having biological activities, useful
PT as nutritional sources or supplements, or for treating e.g. autoimmune
PT diseases, cancers, bone fractures or damages, or central nervous system
PT disorders.
XX
PS Claim 12; Page 109; 288pp; English.
XX
CC The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides can be used to express recombinant proteins
CC for analysis, characterisation or therapeutic use, as markers for tissues
CC in which the corresponding protein is expressed, as molecular weight
CC markers on Southern gels, as chromosome markers or tags to identify

```

CC chromosomes or to map related gene positions, to compare with endogenous
CC DNA sequences in patients to identify potential genetic disorders, as
CC probes to hybridise and discover novel related DNA sequences, as a source
CC of information to derive PCR primers for genetic fingerprinting, to raise
CC anti-protein antibodies and in gene therapy. The proteins can be used to
CC raise antibodies or to elicit another immune response, as reagents in
CC assays designed to quantitatively determine levels of the protein in
CC biological fluids, as markers for tissues in which the corresponding
CC protein is preferentially expressed and to treat autoimmune disorders
CC (e.g. multiple sclerosis, systemic lupus erythematosus, insulin dependent
CC diabetes mellitus or graft-versus-host disease), anaemia, periodontal
CC diseases, bone fractures, cartilage damage, central nervous system
CC disorders (e.g. Alzheimer's disease or Parkinson's disease) and cancers.
CC The proteins and polynucleotides are also useful as nutritional sources
CC or supplements, e.g. as carbon, nitrogen or carbohydrate sources. This
CC sequence represents a human polynucleotide of the invention.

XX Sequence 1238 BP, 387 A, 336 C, 266 G, 249 T, 0 U, 0 Other;

Query Match 41.5%; Score 1008.2; DB 9; Length 1238;

Best Local Similarity 97.9%; Pred. No. 3.5e-220; Mismatches 3; Indels 23; Gaps 19;

Matches 1223; Conservative 0; Mismatches 3; Indels 23; Gaps 19;

QY 1181 GAGCAGCTCATCAACCCCTTTGAGAGAGATGATGATTTTGAAGCACTGATTTGTC 1240
DB 1 GAGCAGCTCATCAACCCCTTTGAGAGAGATGATGATTTTGAAGCACTGATTTGTC 60
QY 1241 GACAGGAATTTGACAGTGTCTCTGTGGCTGTGATGATGATTCACAGAGCTGCTGG 1300
DB 61 GACAGGAATTTGACAGTGTCTCTGTGGCTGTGATGATGATTCACAGAGCTGCTG 119
QY 1301 ATGGAGCCGAGCATGTATCTGAAATPAGCCGAGCAGAGCCGCCCTTACAGAGCTGCTCC 1360
DB 120 ATGGAGCCGAGCATGTATCTGAAATPAGCCGAGCAGAGCCGCCCTTACAGAGCTGCTC 178
QY 1361 GCCAGTTCGCTGAGACCTCTTTATGGGCTTCCACCTTCAACATGAGCTTGAACAGAG 1420
DB 179 GCCAGTTCGCTGAGACCTCTTTATGGGCTTCCACCTTCAACATGAGCTTGAACAGAG 237
QY 1421 GAGATGAGTTCAGAGCCCAATCAGAGAGACGAGAGATGCTCAGCTGAGATCATTTGG 1480
DB 238 GAGATGAGTTCAGAGCCCAATCAGAGAGACGAGAGATGCTCAGCTGAGATCATTTGG 296
QY 1481 GCGTTCCAGAGGCTGAGTCCCATATGATCAACATCTCCAGAGGCAACCTCAAGAGCAAA 1540
DB 297 GCGTTCCAGAGGCTGAGTCCCATATGATCAACATCTCCAGAGGCAACCTCAAGAGCA 355
QY 1541 CTACTGTGGCCCAAGAGGAAATCCCTTCTCAGAGAGGCTGCCCCAAAACACAGAGCA 1600
DB 356 CTACTGTGGCCCAAGAGGAAATCCCTTCTCAGAGAGGCTGCCCCAAAACACAGAGGC 414
QY 1601 GCCAAACAGAACTTTAGGGGCCAGAGACAAAGAGCTTGAAGCTTGAAGCTGTGGAC 1660
DB 415 GCCAAACAGAACTTTAGGGGCCAGAGACAAAGAGCTTGAAGCTTGAAGCTGTGGAC 473
QY 1661 GCCTTCAAGTCTGGGCCAATGATCAGAGGCGAGCTCTCAAGTGGCCCAAGAGCGCC 1720
DB 474 GCCTTCAAGTCTGGGCCAATGATCAGAGGCGAGCTCTCAAGTGGCCCAAGAGCGCC 532
QY 1721 CTCAGCCCCACTCCCATGTTCTTCCCTTAGAACCATGAGCGCTCAAGCTTCAACAGT 1780
DB 533 CTCAGCCCCACTCCCATGTTCTTCCCTTAGAACCATGAGCGCTCAAGCTTCAACAG 591
QY 1781 GTCACAGGATAGACACCAAGACAAAGCTTAAAGACTGTGAGTGTCTGGGGCCAGAAA 1840
DB 592 GTCACAGGATAGACACCAAGACAAAGCTTAAAGACTGTGAGTGTCTGGGGCCAGAA 650
QY 1841 AGTTTGAATGCTCTCAGAGAGGATGGGGCTTGAATGAGACCCAGAAAGTATCTCAA 1900
DB 651 AGTTTGAATGCTCTCAGAGAGGATGGGGCTTGAATGAGACCCAGAAAGTATCTC-A 709
QY 1901 GTGAGAGAGAAACTGTGAGATTTAACTGACGATATGCGAGAGATCCCGAAATATCAC 1960

DB 710 GTGAGAGAGAAACTGTGAGATTTAACTGACGATATGCGAGAGATCCCGAAATATCA- 768
QY 1961 CTCAAAGAACCTTTGGAAACAATCAACCAACCAATATACATCACTACATCAAGATGACATG 2020
DB 769 CTCAAAGAACCTTTGGAAACAATCAACCAACCAATATACATCACTACATCAAGATATCAT- 827
QY 2021 GATCCTTATTTGGGCTTTGGAAACAGGGATGAGACATCTCTAACCTGCTTCTTAATGG 2080
DB 828 GATCCTTATTTGGGCTTTGGAAACAGGGATGAGACATCTCTAACCTGCTTCTTAATG 886
QY 2081 GGAATGCTTGGCAGCAGAGTCTCTCACTGTGTGTACACAGAGACACTGATCCAGTCA 2140
DB 887 GGAATGCTTGGCAGCAGAGTCTCTCACTGTGTGTACACAGAGACACTGATCCAGTC- 945
QY 2141 CAGCATACAGCTGTGCCACATGAAACGAGTCTTCAAGAGCCTGAATCAAAATGTTA 2200
DB 946 CAGCATACAGCTGTGCCACATGAAACGAGTCTTCAAGAGCCTGAATCAAAATG--C 1003
QY 2201 GCTTAAATGATTAATAATCCAGACTACTTCAAGCCTTTAATGCTTTAATTAATAAACT 2260
DB 1004 GCTTAAATGATTAATAATCCAGACTACTTCAAGCCTTTAATGCTTTAATTAATAAAA-- 1061
QY 2261 GTGAAAGCTTACATGAAACCATTTGAAACATTTAACTCAGACTCTGAGATTCAAGTCGGA 2320
DB 1062 GTGAAAGCTTACATGAAACCATTTGAAACATTTAACTCAGACTCTGAGATTCAAGTCG 1119
QY 2321 ACCCTTGTCTATCTGAAATCCAGAGAGCAGCAACCTTAGTACTGCGCAAACTAATGA 2380
DB 1120 ACCCTTGTCTATCTGAAATCCAGAGAGCAGCAACCTTAGTACTGCGCAAACTAAT-- 1177
QY 2381 GTTTAATTAATTAATAAATCTGTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 2429
DB 1178 GTTTAATTAATTAATAAATCTGTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 1226

RESULT 8
AAZ21229
ID AAZ21229 standard, cDNA, 1916 BP.
XX
AC AAZ21229;
XX
DT 22-NOV-1999 (first entry)
XX
DE Mouse CG1CE cDNA sequence.
XX
XX CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
XX age-related macular dystrophy; ss.
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 11..166
FT /tag= a
FT /product= "CG1CE protein"
XX
PN W09943695-A1.
XX
PD 02-SEP-1999.
XX
PF 22-FEB-1999; 99WO-US003790.
XX
PR 25-FEB-1998; 98US-0075941P.
PR 18-DEC-1998; 98US-0112926P.
XX
PA (MERI) MERCK & CO INC.
PA (UNIP-) UNIV UPSALA.
XX
PI Petrunkin K, Caskey CT, Metzker M, Madeline C;
XX WPI; 1999-540560/45.
DR P-PSDB; AAY29955.
XX
XX Human and mouse polynucleotides encoding CG1CE polypeptides.

XX Claim 2; Fig 8; 67bp; English.

PS The present sequence represents the mouse CGICE cDNA sequence, which when
CC mutated is responsible for Best's macular dystrophy (BMD).
CC Polynucleotides encoding CGICE are useful for diagnosing whether a
CC patient carries a mutation in the CGICE gene. Normal and mutated CGICE
CC proteins are useful for identifying activators and/or inhibitors of these
CC proteins, in order to treat BMD. The CGICE gene offers a simpler and
CC cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
XX age-related macular dystrophy

SQ Sequence 1916 BP; 514 A; 499 C; 457 G; 446 T; 0 U; 0 Other;

Query Match 27.1%; Score 658.2; DB 2; Length 1916;
Best Local Similarity 67.6%; Pred. No. 4.1e-140;
Matches 1116; Conservative 0; Mismatches 293; Indels 242; Gaps 4;

```

QY 96 TGCCTGCGCATGACATCACTTACAGAGCCAAAGTGGCTAATGCCGCTTAGGCTCTTC 155
DB 2 TGCACAGCCATGACTATCACTTACACAAAGTGGCCAAATGCCGCTCGGTTCTTC 61
QY 156 TCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 215
DB 62 TCGTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
QY 216 ATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 275
DB 122 GTCTTATATTCCTCTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 181
QY 276 CAAAGCTGATGTTTGAAGAACTGACTCTGTAATGCGACAGCTAGATCACTGCTCC 335
DB 182 CAGCACTGTTGTTGAAGAGCTGGCTGTACTGGAAGCTAGATCACTGCTGCTGCTGCT 241
QY 336 ATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 395
DB 242 ATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
QY 396 GAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
DB 302 GAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
QY 456 GACGAGCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 515
DB 362 GATGGAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
QY 516 CTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 575
DB 422 CTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
QY 576 GTGCAAGAGCTTATATGACTTCGCGAGAACAAAGAGTTGAGAACTGAGCTTACA 635
DB 482 GTGCAAGAGCTTATATGACTTCGCGAGAACAAAGAGTTGAGAACTGAGCTTACA 541
QY 636 CACAACATGTTCTGGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 695
DB 542 CACAACATGTTCTGGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
QY 696 GAGAGTGAATCCGGAACCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 755
DB 602 GAGAGTGAATCCGGAACCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
QY 756 GGTACTGAGTGTGACACCTGTATGCTTACGATGTAATGATCCCACTGCTGTATACA 815
DB 662 GGTACTGAGTGTGACACCTGTATGCTTACGATGTAATGATCCCACTGCTGTATACA 721
QY 816 CAGGTGTGACTGTGCGGTGTACAGCTTCTTCTGACTGTTGAGTGGCGGCTGTT 875
DB 722 CAGGTGTGACTGTGCGGTGTACAGCTTCTTCTGACTGTTGAGTGGCGGCTGTT 781
QY 876 CTGAACCCAGCAAGGCTTACCTTGGCCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCT 935
DB 876 CTGAACCCAGCAAGGCTTACCTTGGCCATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCT
```

```

DB 782 CTGAACCCAAACAAAGACTACCCAGGCCATGATGATCTGTGTGCTGTCTTACA 841
QY 936 TTCTGAGATCTTCTTCTTATGTTTGGCTGCTGAGAGTGGGCTCTCCAGGGCCCTG 995
DB 842 ATCTGCAATCTTATTTTCAATGAGCTGGGCTGAA----- 876
QY 996 GGCTGAGGCAATGGCCAGAGGGGTTCATGGCCAGCAAGCTGCTTGAAGCAGATGACGTGT 1055
DB 877 ----- 876
QY 1056 CAGAAAGAAAGTCTACAGGGGTAGAAAGCAGCCAGCGCTGTGTGGCCACACTGTATC 1115
DB 877 ----- 876
QY 1116 CCACTACTCGGAGGCTGAGGCAAGAAATCGCTTGAACCCGAGGCGGAGGTTGTG 1175
DB 877 -----GG 878
QY 1176 TGGGAGGCAAGCTCATACACCCCTTTGAGAGATGATGATTTTGAAGCAACTGGA 1235
DB 879 TGGGAGGCAAGCTCATACACCCCTTTGAGAGATGATGATTTTGAAGCAACTGGA 938
QY 1236 TTGTGCAAGAAATTTGACAGTGTCCCTGTTGGCTGTGATGATGATGATGATGATGATGAT 1295
DB 939 TCATTTGACAGAACTGTGACAGTGTCCCTGTTGGCTGTGATGATGATGATGATGATGATGAT 998
QY 1296 CTGCGATGAGCGGACATGTACTGTGAATTAAGCCGAGCCAGCCGCCCTTACACAGCTG 1355
DB 999 CTCCATGAGAACTGATCATGTACTGTGAACGAGGCAAGCGCTGAGCGGCTTACAGAGCTG 1058
QY 1356 CTTCGCGGCTTCCGTCAGGCTCTCTTTATGAGGCTTCCACCTTGAACATCAGCTTGAACA 1415
DB 1059 CTTCGCGGCTTCCGTCAGGCTCTCTTTATGAGGCTTCCACCTTGAACATCAGCTTGAACA 1118
QY 1416 AAGAGATGAGATTTCAGCCCAATCAGAGGAGCAGAGATGCTCAGCTG----- 1469
DB 1119 AAGAGATGAGATTTCAGCCCAATCAGAGGAGCAGAGATGCTCAGCTG----- 1178
QY 1470 -----GCATCATTTGGCCGCTTCTGAGGCTGCTGATGATGATGATGATGATGATGATGAT 1523
DB 1179 ATGAGCAACCAATGAGCTCTCTTTTGAAGCTCAACCCAAATCAATCAATCTTCCCTTGA 1238
QY 1524 CAAACTCAAGCAACCAATGAGCTCTCTTTTGAAGCTCAACCCAAATCAATCAATCTTCCCTTGA 1583
DB 1239 AAGACTTAAAGCAACCAATGAGCTCTCTTTTGAAGCTCAACCCAAATCAATCAATCTTCCCTTGA 1292
QY 1584 CCAAAAACCAAGGCAAGCCAAACAGAACCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1643
DB 1293 GTAAAGATGCCAAACCAAAAACAGAAAG-----ATGCTGGA 1331
QY 1644 AGCTTAAGCTGTGAGACGCTTCAAGTCTGAGCCCATGATGAGAGGCAAGCTTAC 1703
DB 1332 AATTTAAGGCTGTGAGCTTCTTGAATGTGTTTCAAGGTTTAAAGGAGAGGCTTCCATT 1391
QY 1704 GTGCCCAAGACGCTTCAAGCCCACTCC 1734
DB 1392 GTGCCCAAGACGCTTCAAGCCCACTCC 1422
RESULT 9
AAZ21226
ID AAZ21226 standard; DNA; 16125 BP.
XX
AC AAZ21226;
XX
DT 22-NOV-1999 (first entry)
XX
XX Human CGICE genomic DNA sequence.
XX
XX CGICE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
XX age-related macular dystrophy; se.
XX
OS Homo sapiens.
```

XX	Key	Location/Qualifiers
PH	Intron	1..1513
FT		/*tag= a
FT		/number= 1
FT	exon	1514..1570
FT		/*tag= b
FT		/number= 1
FT	Intron	1571..3278
FT		/*tag= c
FT		/number= 2
FT	exon	3279..3466
FT		/*tag= d
FT		/number= 2
FT	CDS	3315..15787
FT		/*tag= e
FT		/product= "CGICE protein"
FT		/note= "contains introns"
FT	Intron	3467..6616
FT		/*tag= f
FT		/number= 3
FT	exon	6617..6711
FT		/*tag= g
FT		/number= 3
FT	Intron	6712..7227
FT		/*tag= h
FT		/number= 4
FT	exon	7228..7461
FT		/*tag= i
FT		/number= 4
FT	Intron	7462..8353
FT		/*tag= j
FT		/number= 5
FT	exon	8354..8508
FT		/*tag= k
FT		/number= 5
FT	Intron	8509..8896
FT		/*tag= l
FT		/number= 6
FT	exon	8897..8974
FT		/*tag= m
FT		/number= 6
FT	Intron	8975..9676
FT		/*tag= n
FT		/number= 7
FT	exon	9677..10032
FT		/*tag= o
FT		/number= 7
FT	Intron	10033..11212
FT		/*tag= p
FT		/number= 8
FT	exon	11213..11293
FT		/*tag= q
FT		/number= 8
FT	Intron	11294..11556
FT		/*tag= r
FT		/number= 9
FT	exon	11557..11708
FT		/*tag= s
FT		/number= 9
FT	Intron	11709..13918
FT		/*tag= t
FT		/number= 10
FT	exon	13919..14557
FT		/*tag= u
FT		/number= 10
FT	Intron	14558..15768
FT		/*tag= v
FT		/number= 11
FT	exon	15769..16125
FT		/*tag= w
FT		/number= 11
XX		

PN	W09943695-A1.
XX	02-SEP-1999.
PD	
XX	
PF	22-FEB-1999; 99WO-US003790.
XX	
PR	25-FEB-1998; 98US-0075941P.
PR	18-DEC-1998; 98US-0112926P.
XX	
PA	(MERI) MERCK & CO INC.
PA	(UTUP-) UNIV UPPSALA.
PI	Petrukhin K, Caskey CT, Metzker M, Wadelius C;
XX	
XX	WPI; 1999-540560/45.
DR	P-PSDB; AAY29953.
XX	
PT	Human and mouse polynucleotides encoding CGICE polypeptides.
XX	
PS	Claim 2; Fig 1; 67pp; English.
XX	
CC	The present sequence represents the human CGICE gene, which when mutated
CC	is responsible for Best's macular dystrophy (BMD). Polynucleotides
CC	encoding CGICE are useful for diagnosing whether a patient carries a
CC	mutation in the CGICE gene. Normal and mutated CGICE proteins are useful
CC	for identifying activators and/or inhibitors of these proteins, in order
CC	to treat BMD. The CGICE gene offers a simpler and cheaper method of
CC	diagnosing BMD without the need for the presence of the patient. The gene
CC	may also be useful to discovering the genetic cause of age-related
CC	macular dystrophy
XX	
SQ	Sequence 16125 BP; 3988 A; 4175 C; 4215 G; 3726 T; 0 U; 21 Other;
Query Match	26.4%; Score 642.2; DB 2; Length 16125;
Best Local Similarity	99.5%; Pred. No. 4e-136;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	1405 CAGCTGAACAAAGAGAGATGAGTTCACCCCAATCAGAGAGAGAGATGCTCA 1464
DB	13916 CAGCTGAACAAAGAGAGATGAGTTCACCCCAATCAGAGAGAGATGCTCA 13975
QY	1465 CGCTGGCATCTTTGGCGCTTCTAGGCTGCACTCCCATGATCACTCTCCAGGGC 1524
DB	13976 CGCTGGCATCTTTGGCGCTTCTAGGCTGCACTCCCATGATCACTCTCCAGGGC 14035
QY	1525 AAACCTCAAGACCAAACTACTGTGCGCCCAAGAGGGAATCCTTCTCCAGAGGCTTCC 1584
DB	14036 AAACCTCAAGACCAAACTACTGTGCGCCCAAGAGGGAATCCTTCTCCAGAGGCTTCC 14095
QY	1585 CAAAACACAGAGCAGCCAAACAGAAAGTTAGGGGCGAGAGACAAAGGCGCTGGAA 1644
DB	14096 CAAAACACAGAGCAGCCAAACAGAAAGTTAGGGGCGAGAGACAAAGGCGCTGGAA 14155
QY	1645 GCTTAAAGCTGTGAGCGCTTCAAGTCTGCGCCCACTGATCAGAGGCGCACTACTACG 1704
DB	14156 GCTTAAAGCTGTGAGCGCTTCAAGTCTGCGCCCACTGATCAGAGGCGCACTACTACG 14215
QY	1705 TGCCCCACAGAGCGCCCTCAGCGCCCACTCCCATGTTCTTCCCTAGAACATCAGCGCC 1764
DB	14216 TGCCCCACAGAGCGCCCTCAGCGCCCACTCCCATGTTCTTCCCTAGAACATCAGCGCC 14275
QY	1765 GTCAAAGCTTCACAGTGCACAGCATAGACACAAAGCAAAAGCTTAAAGACTGTGAG 1824
DB	14276 GTCAAAGCTTCACAGTGCACAGCATAGACACAAAGCAAAAGCTTAAAGACTGTGAG 14335
QY	1825 TTCTGGGGCCAAAGAAATTGTAATGCTCTCAGAGAGCGATGGGCTTTAGTAGACA 1884
DB	14336 TTCTGGGGCCAAAGAAATTGTAATGCTCTCAGAGAGCGATGGGCTTTAGTAGACA 14395
QY	1885 CCCAAGATATCTCAAGTGAAGAGAAACTGTGGAAGTTTAACTGACGATATGCGACA 1944
DB	14396 CCCAAGATATCTCAAGTGAAGAGAAACTGTGGAAGTTTAACTGACGATATGCGACA 14455

QY 1945 GATCCCGAAATCCTCAAGAACCTTGGACATCCACCAACATACACTAC 2004
DB 14456 GATCCCCGAAAAATCCTCAAGAACTTGGAAACATACCAACCACTACACTAC 14515
QY 2005 ACTCAAGATCATGATGCTTATGGGCTTGGAAACAGGAG 2051
DB 14516 ACTCAAGATCATGATGCTTATGGGCTTGGAAACAGGCTCG 14562

RESULT 10
ABAI4559
ID ABAI4559 standard; DNA; 16650 BP.
AC ABAI4559;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6890.
XX
KW Human; nootropic; neuroprotective; cyostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antisticking; antianaemic; antiarthritic; cancer;
KW antidiabetic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 02-MAR-2000; 2000US-0184664P.
PR 16-MAR-2000; 2000US-0186350P.
PR 17-MAR-2000; 2000US-0189874P.
PR 18-APR-2000; 2000US-0190076P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 14-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.

PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 27-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.

PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246533P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMAN -) HUMAN GENOME SCI INC.
PA
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX
PS Disclosure; SEQ ID NO 6888; 1701bp + Sequence listing; English.
XX
CC The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABA1678-ABA18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease.

PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0242221P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246539P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0251799P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.

PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCT INC.
XX
XX Rosen CA, Barash SC, Ruben SM,
XX WPI, 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX
XX
XX Disclosure; SEQ ID NO 6889; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
CC (AB11478-AB11801) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at fcp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 18537 BP; 4519 A; 4800 C; 4983 G; 4235 T; 0 U; 0 Other;
Query Match 26.4%; Score 640.6; DB 5; Length 18537;
Best Local Similarity 99.4%; Pred. No. 9,8e-136;
Matches 643; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1405 CAGCTGAACCAAGAGAGATGAGTTCAGCCCAATCAGAGAGACGAGAGATGCTCA 1464
DB 11901 CAGCTGAACCAAGAGAGATGAGTTCAGCCCAATCAGAGAGACGAGATGCTCA 11960
QY 1465 CGCTGGCATCTTGGCCGCTTCCTAGGCTGCAAGTCCCATATCAATCTCTCCAGGCG 1524
DB 11961 CGCTGGCATCTTGGCCGCTTCCTAGGCTGCAAGTCCCATATCAATCTCTCCAGGCG 12020
QY 1525 AACTCAAGACCAAACTACTGTGGCCCAAGAGGAAATCCCTTCCACGAGGCGCTGGC 1584
DB 12021 AACTCAAGACCAAACTACTGTGGCCCAAGAGGAAATCCCTTCCACGAGGCGCTGGC 12080
QY 1585 CAAAACCAAGAGGAGCCCAAGAAAGCTTAGGGCCAGAGAGACAAAGGCTTGGAA 1644
DB 12081 CAAAACCAAGAGGAGCCCAAGAAAGCTTAGGGCCAGAGAGACAAAGGCTTGGAA 12140
QY 1645 GCTTAAGGCTGTGACGCTTCAAGTGTGGCCCACTGTATCAGAGGCGAGCTACTACG 1704
DB 12141 GCTTAAGGCTGTGACGCTTCAAGTGTGGCCCACTGTATCAGAGGCGAGCTACTACG 12200
QY 1705 TGCCCAAGAGCGCCCTCAGCCCACTCCCATGTTCTTCCCTTAGAACCATCAGCGCC 1764
DB 12201 TGCCCAAGAGCGCCCTCAGCCCACTCCCATGTTCTTCCCTTAGAACCATCAGCGCC 12260
QY 1765 GTCAAGCTTCAAGTGTCAAGGATAGACCAAGAAAGCTTAAAGCTGTGAG 1824
DB 12261 GTCAAGCTTCAAGTGTCAAGGATAGACCAAGAAAGCTTAAAGCTGTGAG 12320
QY 1825 TTCTGGGGCCCAAGAAAGTTTAAATTGCTCTCAGAGACGATGGGGCTTATGAGACA 1884
DB 12321 TTCTGGGGCCCAAGAAAGTTTAAATTGCTCTCAGAGACGATGGGGCTTATGAGACA 12380
QY 1885 CCAGAAATCTCAAGTGTGAGAGAAAGCTGTGAGTTAATCTGACGATATGCCAGA 1944

Db 12381 CCCAAGACTATCTCAAGTAGAGAGAAAACTGTGAGTTTAACTGACGATATGCCAGA 12440
Qy 1945 GATCCCGGAAATACCTCAAGAAACCTTTGGACATACCAACCAACATACACACTAC 2004
Db 12441 GATCCCGGAAATACCTCAAGAAACCTTTGGACATACCAACCAACATACACACTAC 12500
Qy 2005 ACTCAAGATCATGATGATCTTTATTTGGGCGCTTGGAAAAACAGGATG 2051
Db 12501 ACTCAAGATCATGATGATCTTTATTTGGGCGCTTGGAAAAACAGGCTGTG 12547

RESULT 13
ABL89697
ID ABL89697 standard; cDNA; 1717 BP.
XX
AC ABL89697;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 259.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antidiabetic; hepatotropic; antidiabetic; antiinflammatory; antiviral;
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PP 18-MAY-2001; 2001WO-US016450.
XX
PR 19-MAY-2000; 2000US-0205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
XX WPI; 2002-122018/16.
DR P-PSDB; ABB89288.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
PS Claim 4; SEQ ID NO 259; 2081bp + Sequence Listing; English.

XX The invention relates to novel genes (ABL89449-ABJ90853) and proteins
XX (ABB89040-ABB89044) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemia; (d) wound healing
XX / (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at http://wipo.int/pub/published_pct_sequences
XX
XX Sequence 1717 BP; 464 A; 460 C; 417 G; 367 T; 0 U; 9 Other;

Query Match 24.2%; Score 587.8; DB 6; Length 1717;
Best Local Similarity 98.8%; Pred. No. 4.9e-124;

Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
Qy 1445 GAGAGCGAGGAGATGTCACGCTGAGCATCTTGGCCGCTTCCTGAGGCTGAGTCCAT 1504
Db 1 GAGAGCGAGGAGATGTCACGCTGAGCATCTTGGCCGCTTCCTGAGGCTGAGTCCAT 60
Qy 1505 GATCAACATCTCTCCAGGGCAAACTCAAGAGCAAACTACTGTGGCCCAAGAGGAATCC 1564
Db 61 GATCAACATCTCTCCAGGGCAAACTCAAGAGCAAACTACTGTGGCCCAAGAGGAATCC 120
Qy 1565 CTCTCCAGAGGGCTGCGCCCAAAACCAAGGACGCAAAACGATTTAGGGGCCAG 1624
Db 121 CTCTCCAGAGGGCTGCGCCCAAAACCAAGGACGCAAAACGATTTAGGGGCCAG 179
Qy 1625 GAAGCAACAAGGGCTGAGAGCTTAAGGCTGTGACGCTTCAAGTCTGGGCCACTGTAT 1684
Db 180 GAAGCAACAAGGGCTGAGAGCTTAAGGCTGTGACGCTTCAAGTCTGGGCCACTGTAT 239
Qy 1685 CAGAGGCCAGGCTACTACAGTGGCCCAAGAGCGCCCTCAGCCCTCCATGTTCTTC 1744
Db 240 CAGAGGCCAGGCTACTACAGTGGCCCAAGAGCGCCCTCAGCCCTCCATGTTCTTC 299
Qy 1745 CCCCTGAACCATCAGCGCCGCTCAAGCTTCAAGTGTCAAGGATAGACACCAAGAC 1804
Db 300 CCCCTGAACCATCAGCGCCGCTCAAGCTTCAAGTGTCAAGGATAGACACCAAGAC 359
Qy 1805 AAAAGCTTAAGACGTGATGTTGGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGC 1864
Db 360 AAAAGCTTAAGACGTGATGTTGGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGC 419
Qy 1865 GATGGGGCTTGAATGAGACACCCAGAAAGTATCTCAAGTGAAGAGAGAAACTGTGAGTTT 1924
Db 420 GATGGGGCTTGAATGAGACACCCAGAAAGTATCTCAAGTGAAGAGAGAAACTGTGAGTTT 479
Qy 1925 AACCTGACGATATGTCAGAGATCCCGAAATATCACTCAAGAAACCTTTGGAACATCA 1984
Db 480 AACCTGACGATATGTCAGAGATCCCGAAATATCACTCAAGAAACCTTTGGAACATCA 539
Qy 1985 CCAACCAACATCACTCAAGTCTCAAGATCAATGATCTTATTTGGGCGCTTGGAAAC 2044
Db 540 CCAACCAACATCACTCAAGTCTCAAGATCAATGATCTTATTTGGGCGCTTGGAAAC 599
Qy 2045 AGGATG 2051
Db 600 AGGCTGTG 606

RESULT 14
ABAI4556
ID ABAI4556 standard; DNA; 7108 BP.
XX
AC ABAI4556;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6887.
XX
KW Human; nootropic; neuroprotective; cyostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antiskinking; antianaemic; antiarthritis; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antidiabetic; antidiabetic; antidiabetic; antidiabetic; antidiabetic;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; neurotropic; gene therapy; vaccine; ds.
XX
XX Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX

PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225470P.
PR 14-AUG-2000; 2000US-0225477P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0233080P.
PR 08-SEP-2000; 2000US-0233081P.
PR 12-SEP-2000; 2000US-0233968P.
PR 14-SEP-2000; 2000US-0233979P.
PR 14-SEP-2000; 2000US-0233989P.
PR 14-SEP-2000; 2000US-0234000P.
PR 14-SEP-2000; 2000US-0234001P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236337P.
PR 29-SEP-2000; 2000US-0236357P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0242221P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251889P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
DR
XX
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT and for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.

XX Disclosure; SEQ ID NO 6887; 1701bp + Sequence Listing; English.
PS
XX The invention relates to novel genes (ABAI1004-AB21334) and proteins
CC (AB21678-AB21801) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 7108 BP; 1632 A; 1882 C; 1989 G; 1605 T; 0 U; 0 Other;
Query Match 19.6%; Score 476.4; DB 5; Length 7108;
Best Local Similarity 99.4%; Pred. No. 2.4e-98;
Matches 499; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 1405 CAGCTGAACAAAGAGATGAGATTCAGCCCATATAGAGACGAGAGATGCTCA 1464
DB 6609 CAGCTGAACAAAGAGATGAGATTCAGCCCATATAGAGACGAGAGATG-TCA 6667
QY 1465 CGCTGGCATCATTTGGCGCTTCTTCTAGGCTGACATGATCATCATCTCCAGGGC 1524
DB 6668 CGCTGGCATCATTTGGCGCTTCTTCTAGGCTGACATGATCATCATCTCCAGGGC 6727
QY 1525 AAACCTAAGACCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCAAGAGGCTGCG 1584
DB 6728 AAACCTAAGACCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCAAGAGGCTGCG 6786
QY 1585 CAAACCAACCAAGGACGACCAACAGAACGTTAGGGGCGGAGAAACAAAGGCTGGAA 1644
DB 6787 CAAACCAACCAAGGACGACCAACAGAACGTTAGGGGCGGAGAAACAAAGGCTGGAA 6846
QY 1645 GCTTAAGGCTGTGAGCGCTTCAAGTGTGGCCACTGATAGAGGCGGCTACTACAG 1704
DB 6847 GCTTAAGGCTGTGAGCGCTTCAAGTGTGGCCACTGATAGAGGCGGCTACTACAG 6906
QY 1705 TGCCCCACAGAGCGCCCTCAGCGCCCACTCCATGTTCTTCCCTAGAACCATCAGCGCC 1764
DB 6907 TGCCCCACAGAGCGCCCTCAGCGCCCACTCCATGTTCTTCCCTAGAACCATCAGCGCC 6966
QY 1765 GTCAAGGCTTCAAGTGTACAGGATAGACACCAAGAACAAAGCTTAAGACTGTGAG 1824
DB 6967 GTCAAGGCTTCAAGTGTACAGGATAGACACCAAGAACAAAGCTTAAGACTGTGAG 7026
QY 1825 TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTGATGAGACA 1884
DB 7027 TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTGATGAGACA 7086
QY 1885 CCCAGAAATATCTCAAGTGAGG 1906
DB 7087 CCCAGAAATATCTCAAGTGAGG 7108
RESULT 15
AB061177
ID AB061177 standard; cDNA; 2112 BP.
XX
AC AB061177;
XX
XX 26-FEB-2003 (first entry)
XX
DE FLJ20132 fls clone COL06441 encoding sequence.

XX Neuroprotective; immunomodulator; cancer; chromosome 19; cytoskeletal;
KW anti-inflammatory; gene therapy; nutritional supplement; wound; burn;
KW ulcer; Alzheimer's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
KW vulnheraty; gene; ss.
OS Homo sapiens.
PN WO200231111-A2.
XX
XX 18-APR-2002.
XX
XX 11-OCT-2001; 2001WO-US027760.
XX
XX 12-OCT-2000; 2000US-00687527.
PR
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
XX WPI; 2002-426278/45.
DR
XX N-PSDB; ABP43933.
XX
PT New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation.
XX
XX Claim 1; SEQ ID # 390; 357bp + Sequence Listing; English.
PS
XX The invention relates to 446 newly isolated polynucleotide sequences. The
CC activity of polynucleotides of the invention may be described as,
CC vulnheraty, neuroprotective, immunomodulator, cytostatic and anti-
CC inflammatory. Compositions comprising nucleic acids of the invention are
CC useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records AB060788-
CC AB061233 represent polynucleotides of the invention. NOTE: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2112 BP; 382 A; 681 C; 607 G; 442 T; 0 U; 0 Other;
Query Match 17.8%; Score 431.4; DB 6; Length 2112;
Best Local Similarity 60.4%; Pred. No. 2.9e-88;
Matches 597; Conservative 0; Mismatches 276; Indels 0; Gaps 0;
QY 103 CCATGACCATCACTTATACACAGCCAGATGAGTATGCGGCTTCTTCCCGCC 162
DB 190 CGATGACCGTCACTTACACAGCCAGATGAGTATGCGGCTTCTTCCCGCC 249
QY 163 TGTGCTGTGTGCGGGGCGGACATCTACAGCTGCTATATGCGGAGTTCTTAATCTTCC 222
DB 250 TGTGCTGTGTGCGGGGCGGACATCTACAGCTGCTATATGCGGAGTTCTTGTCTTCC 309
QY 223 TGTGCTGTGTGCGGGGCGGACATCTTATATGAGCTGCGGCTTACGGAAGAACACAGC 282
DB 310 TTGGGTTCTACATGCGCGCTGAGTGTCTCTTGTGCTTACGGAAGGAGAGAGC 369
QY 283 TGATGTTTGAAACTGACTCTGTATTTGCGACAGCTACATCAAGCTATCCCATTTTCT 342
DB 370 GCTACTTTCGAGAGCTGTGATTTATTTGACACAGTATGCGAGCTCATCCCTGTCTCCT 429
QY 343 TGTGCTGTGCGCTTACGTGACGCTGTGCTGTGACCGCTGTGGAACCAAGTACGAGAAC 402
DB 430 TGTGCTGTGCGCTTATGATGACGCTGTGGAACCGCTGTGGAACCAAGTACCTATGCA 489

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 26, 2005, 07:39:17 ; Search time 1353.06 Seconds
(without alignments)
10698.156 Million cell updates/sec

Title: US-09-622-964A-4

Perfect score: 2429

Sequence: 1 caggagagctccaccagccta.....aaaaaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5552208 seqs, 297965951 residues

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1541.8	63.5	1758	US-10-492-032-1	Sequence 1, Appl1
2	1244.2	51.2	1263	US-09-746-783-3	Sequence 3, Appl1
3	587.8	24.2	1717	US-10-264-237-259	Sequence 259, App
4	431.4	17.8	2137	US-10-492-032-3	Sequence 3, Appl1
5	431	17.7	1530	US-10-492-032-30	Sequence 30, Appl
6	350.4	14.4	1326	US-10-071-766-10	Sequence 7, Appl1
7	339.4	14.0	1422	US-10-492-032-7	Sequence 1, Appl1
8	338.2	13.9	2500	US-10-108-260A-1299	Sequence 1299, Ap
9	336.8	13.9	2028	US-10-492-032-5	Sequence 5, Appl1
10	213	8.8	620	US-09-814-353-19504	Sequence 19504, A
11	189.2	7.8	1292	US-10-198-846-11070	Sequence 11070, A

C 12	175	7.2	1198	9	US-09-880-107-2174	Sequence 2174, Ap
C 13	161.4	6.6	578	13	US-10-027-632-134530	Sequence 134530,
C 14	161.4	6.6	578	17	US-10-027-632-134530	Sequence 134530,
C 15	149.4	6.2	748	13	US-10-027-632-134529	Sequence 134529,
C 16	149.4	6.2	748	17	US-10-027-632-134529	Sequence 134529,
C 17	135.2	5.6	539	16	US-10-029-386-7583	Sequence 7583, Ap
C 18	134.2	5.5	235	16	US-10-029-386-7583	Sequence 21283, A
C 19	126.8	5.2	1350	9	US-09-768-826-16	Sequence 16, Appl
C 20	126.8	5.2	1350	18	US-10-874-484-16	Sequence 16, Appl
C 21	124.4	5.1	2914	18	US-10-723-860-7441	Sequence 7441, Ap
C 22	124.4	5.1	2914	18	US-10-723-860-8303	Sequence 8303, Ap
C 23	116	4.8	853	14	US-10-198-846-11346	Sequence 11346, A
C 24	115.4	4.8	1954	14	US-10-723-860-7576	Sequence 7576, Ap
C 25	104	4.3	462	14	US-10-198-846-2561	Sequence 2561, Ap
C 26	103.6	4.3	3186778	13	US-10-027-632-174961	Sequence 174961,
C 27	103.6	4.3	3186778	17	US-10-027-632-174961	Sequence 174961,
C 28	102	4.2	615	14	US-10-198-846-7843	Sequence 7843, Ap
C 29	98	4.0	507	10	US-09-764-872-195	Sequence 195, Ap
C 30	91	3.7	106707	18	US-10-694-685-8	Sequence 8, Appl1
C 31	90	3.7	624	13	US-10-027-632-243509	Sequence 243509,
C 32	90	3.7	624	17	US-10-027-632-243509	Sequence 243509,
C 33	90	3.7	267156	11	US-09-868-007A-427	Sequence 427, App
C 34	90	3.7	684973	9	US-09-263-959-1	Sequence 1, Appl1
C 35	89.4	3.7	398287	18	US-10-741-601-5719	Sequence 5719, Ap
C 36	89.4	3.7	398287	19	US-10-741-601-17839	Sequence 17839, A
C 37	89	3.7	31994	9	US-09-764-860-599	Sequence 599, App
C 38	89	3.7	31994	9	US-09-764-904-71	Sequence 71, Appl
C 39	89	3.7	31994	14	US-10-091-548-71	Sequence 599, App
C 40	89	3.7	31994	17	US-10-074-095-599	Sequence 599, App
C 41	89	3.7	31994	14	US-10-212-872-559	Sequence 599, App
C 42	88.4	3.6	723	13	US-10-027-632-11024	Sequence 11024, A
C 43	88.4	3.6	723	17	US-10-027-632-11024	Sequence 11024, A
C 44	88.4	3.6	727	13	US-10-027-632-148248	Sequence 148248,
C 45	88.4	3.6	727	13	US-10-027-632-148249	Sequence 148249,

ALIGNMENTS

RESULT 1
US-10-492-032-1
Sequence 1, Application US/10492032
Publication No. US20050049212A1
GENERAL INFORMATION:
APPLICANT: Steuermagel, Arnd
APPLICANT: Brioner, Gunter
APPLICANT: Fritsch, Rudiger
APPLICANT: Eulenberg, Karsten
TITLE OR INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
Regulation of Energy Homeostasis
FILE REFERENCE: 2923-606
CURRENT APPLICATION NUMBER: US/10/492,032
CURRENT FILING DATE: 2004-04-08
PRIOR APPLICATION NUMBER: PCT/EP02/11321
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1758
TYPE: DNA
ORGANISM: Homo sapiens
US-10-492-032-1
Query Match 63.5%; Score 1541.8; DB 19; Length 1758;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 1756; Conservative 0; Mismatches 2; Indels 203; Gaps 1;
Qy 105 ATGACATCACTTACACAGCAAGTGGCTAATGCCGCTTTCCTCCGCTG 164
Db 1 ATGACATCACTTACACAGCAAGTGGCTAATGCCGCTTTCCTCCGCTG 60

QY 165 CTGCTGCTGCGCGGAGAGATCTACAGCTGCTATATGCGAGTTCCTTAATCTTCTG 224
 Db 61 CTGCTGCTGCGCGGAGAGATCTACAGCTGCTATATGCGAGTTCCTTAATCTTCTG 120
 QY 225 CTCTGCTATCATCATTCCTTTATTTATAGCTGCGCCCTCAACGGAATAACAAGCTG 284
 Db 121 CTCTGCTATCATCATTCCTTTATTTATAGCTGCGCCCTCAACGGAATAACAAGCTG 180
 QY 285 ATGTTTGAAGAACTGATCTGTATTTGCGACAGCTTACATCACTCACTCACTTCTTC 344
 Db 181 ATGTTTGAAGAACTGATCTGTATTTGCGACAGCTTACATCACTCACTCACTTCTTC 240
 QY 345 GTGCTGCGCTTCTATCTGATGCTGATGCGCGCTGATGCGCGCTGATGCGCGCTG 404
 Db 241 GTGCTGCGCTTCTATCTGATGCTGATGCGCGCTGATGCGCGCTGATGCGCGCTG 300
 QY 405 CCGTGGCCCGACCGCTCATAGAGCTGTGTGGCTTGTGGAAGCGAAGCAAGCA 464
 Db 301 CCGTGGCCCGACCGCTCATAGAGCTGTGTGGCTTGTGGAAGCGAAGCAAGCA 360
 QY 465 GCGCGGCTGCTGCGCGGAGAGCTCATGCGCTACGCGCAACGCGCAACGCTGCTG 524
 Db 361 GCGCGGCTGCTGCGCGGAGAGCTCATGCGCTACGCGCAACGCGCAACGCTGCTG 420
 QY 525 CCGACGCTACGACCGCGAGCTCATAGAGCTTCCCGACGCGCAACGCTGCTGCAAGCA 584
 Db 421 CCGACGCTACGACCGCGAGCTCATAGAGCTTCCCGACGCGCAACGCTGCTGCAAGCA 480
 QY 585 GCGTTTATGATCTCGCGGAGAACACAGAGCTGAGAGAACTGAGCTTACCAACAGCATG 644
 Db 481 GCGTTTATGATCTCGCGGAGAACACAGAGCTGAGAGAACTGAGCTTACCAACAGCATG 540
 QY 645 TTCTGCGGCTGCGCGGAGAGCTTGTGCGAAGCTGCTCATAGAGCGCTGAGAGCTGCA 704
 Db 541 TTCTGCGGCTGCGCGGAGAGCTTGTGCGAAGCTGCTCATAGAGCGCTGAGAGCTGCA 600
 QY 705 ATCCGGAACCTTATCTGCTGCTCAGAGGCTGCTGAGAGATGAACCTTGTGCTGAG 764
 Db 601 ATCCGGAACCTTATCTGCTGCTCAGAGGCTGCTGAGAGATGAACCTTGTGCTGAG 660
 QY 765 TGTGACACCTGATATGCTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 824
 Db 661 TGTGACACCTGATATGCTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 720
 QY 825 ACTGTGCGGCTGATGAGCTTCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 884
 Db 721 ACTGTGCGGCTGATGAGCTTCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 780
 QY 885 GCCAAGGCTTACCTGCGGAGATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 944
 Db 781 GCCAAGGCTTACCTGCGGAGATGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 QY 945 TTCTTCTTATATGTTGAGCTGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGAG 1004
 Db 841 TTCTTCTTATATGTTGAGCTGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGAG 866
 QY 1005 CATGCGCAGAGGAGTCAATGCGCAGAGCTGCTGAGAGAGATGCTGAGAGAG 1064
 Db 867 CATGCGCAGAGGAGTCAATGCGCAGAGCTGCTGAGAGAGATGCTGAGAGAG 866
 QY 1065 AAGGCTTCAAGGAGTGAAGAGCAGCGGCTGCTGAGAGAGCTGATTCAGAGCTACT 1124
 Db 867 AAGGCTTCAAGGAGTGAAGAGCAGCGGCTGCTGAGAGAGCTGATTCAGAGCTACT 866
 QY 1125 CGGAGGCTGAGGAGAGATGCTTGAACCGGAGAGCGAGGTTGTGTGAGAGC 1184
 Db 867 CGGAGGCTGAGGAGAGATGCTTGAACCGGAGAGCGAGGTTGTGTGAGAGC 877
 QY 1185 AGCTGATCAACCCCTTTGAGAGAGATGATGATTTTGAACAAGCTGATTTGCTGACA 1244
 Db 878 AGCTGATCAACCCCTTTGAGAGAGATGATGATTTTGAACAAGCTGATTTGCTGACA 937

QY 1245 GGAATTTGACAGTGTCTCTGTTGCTGATGATGATGATGATGATGATGATGATG 1304
 Db 938 GGAATTTGACAGTGTCTCTGTTGCTGATGATGATGATGATGATGATGATGATG 997
 QY 1305 AGCGGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1364
 Db 998 AGCGGACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1057
 QY 1365 AGTTCCGTGAGCCCTCTTATGAGCTTCACTTCAATCACTGATGATGATGATG 1424
 Db 1058 AGTTCCGTGAGCCCTCTTATGAGCTTCACTTCAATCACTGATGATGATGATG 1117
 QY 1425 TGAAGTTCAGGCGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1484
 Db 1118 TGAAGTTCAGGCGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1177
 QY 1485 TCTTATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1544
 Db 1178 TCTTATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1237
 QY 1545 TGTGAGCCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1604
 Db 1238 TGTGAGCCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1297
 QY 1605 AACGAACTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1664
 Db 1298 AACGAACTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
 QY 1665 TCAAGTTCAGGCGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1724
 Db 1358 TCAAGTTCAGGCGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1417
 QY 1725 GCGCCATCCCATGTTCTTCCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1784
 Db 1418 GCGCCATCCCATGTTCTTCCCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1477
 QY 1785 CAGGCAATGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1844
 Db 1478 CAGGCAATGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1537
 QY 1845 TTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1904
 Db 1538 TTGAATGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1597
 QY 1905 GAGGAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1964
 Db 1598 GAGGAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1657
 QY 1965 AAGAACCCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 2024
 Db 1658 AAGAACCCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 1717
 QY 2025 CTTATGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2085
 Db 1718 CTTATGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1758

RESULT 2
 US-09-746-783-3
 / Sequence 3, Application US/09746783
 / Publication No. US20030044935A1
 / GENERAL INFORMATION:
 / APPLICANT: Jacobs, Kenneth
 / McCoy, John M.
 / Lavallee, Edward R.
 / Racie, Lisa A.
 / Treacy, Maurice
 / Spaulding, Vikki
 / Agostino, Michael J.
 / Howes, Steven H.
 / Fechele, Kim
 / TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES
 / ENCODING THEM


```
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1596)..(1596)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1634)..(1634)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1657)..(1657)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1702)..(1702)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1712)..(1712)
/ OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-259
```

```
Query Match 24.2%; Score 587.8; DB 17; Length 1717;
Best Local Similarity 98.8%; Pred. No. 1.0e-156;
Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
```

```
QY 1445 GAGACGAGAGAGTGTCTACGCTGGCATCTTGGCCGCTTCTAGAGCCCTGACGTCCCAT 1504
Db 1 GAGACGAGAGAGTGTCTACGCTGGCATCTTGGCCGCTTCTAGAGCCCTGACGTCCCAT 60
QY 1505 GATCACCATCTCTCCAGGGGAAACTCAAGAACCAACTCTGTGGCCCAAGGGAAATCC 1564
Db 61 GATCACCATCTCTCCAGGGGAAACTCAAGAACCAACTCTGTGGCCCAAGGGAAATCC 120
QY 1565 CTCTCTCAGAGGGGCTTGCCTCCAAACCAAGAGCCCAAGAACCAAGAGTTAGGGGGCAG 1624
Db 121 CTCTCTCAGAGGGGCTTGCCTCCAAACCAAGAGCCCAAGAACCAAGAGTTAGGGGGCAG 179
QY 1625 GAAGACAAAGGCTGGAAGCTTAAGGTGTGACGCTTCAAGTCTTGGCCCACTGTAT 1684
Db 180 GAAGACAAAGGCTGGAAGCTTAAGGTGTGACGCTTCAAGTCTTGGCCCACTGTAT 239
QY 1685 CAGAGGCGAGGCTACTACAGTGCCTCCACAGAGCCCTCAGCCCACTCCCATGTTCTTC 1744
Db 240 CAGAGGCGAGGCTACTACAGTGCCTCCACAGAGCCCTCAGCCCACTCCCATGTTCTTC 299
QY 1745 CCCCTAGAACCTACAGGCGCCCTCAAGGCTTCAAGTGTACAGGCAATAGACCAAAAGAC 1804
Db 300 CCCCTAGAACCTACAGGCGCCCTCAAGGCTTCAAGTGTACAGGCAATAGACCAAAAGAC 359
QY 1805 AAAAGCTTAAAGCTGTAGTCTTGGGGGCAAGAAAGTTTGAATGTCTCAGAGAGC 1864
Db 360 AAAAGCTTAAAGCTGTAGTCTTGGGGGCAAGAAAGTTTGAATGTCTCAGAGAGC 419
QY 1865 GATGGGGCTTGTAGTGAACACCCAGAAAGTATCTCAAGTGAAGAGAAACCTGTGAGTTT 1924
Db 420 GATGGGGCTTGTAGTGAACACCCAGAAAGTATCTCAAGTGAAGAGAAACCTGTGAGTTT 479
QY 1925 AACCTGAGATATGCGAGAGATCCCGAAATCAACTCAAAAGCTTGTGAACATCA 1984
Db 480 AACCTGAGATATGCGAGAGATCCCGAAATCAACTCAAAAGCTTGTGAACATCA 539
QY 1985 CCAACCAACATACACATCACTCAAGATACATGATCTTATTTGGGCTTGAAGAAC 2044
Db 540 CCAACCAACATACACATCACTCAAGATACATGATCTTATTTGGGCTTGAAGAAC 599
QY 2045 AGGAGATG 2051
Db 600 AGGATCTG 606
```

RESULT 4
US-10-492-032-3

```
/ Sequence 3, Application US/10492032
/ Publication No. US20050049212A1
/ GENERAL INFORMATION:
/ APPLICANT: Steuermagel, Arnd
/ APPLICANT: Bronner, Gunter
/ APPLICANT: Fritsch, Rudiger
/ APPLICANT: Eulenberger, Karsten
/ APPLICANT: Closssek, Thomas
/ TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
/ TITLE OF INVENTION: Regulation of Energy Homeostasis
/ FILE REFERENCE: 2923-606
/ CURRENT APPLICATION NUMBER: US/10/492,032
/ PRIOR FILING DATE: 2004-04-08
/ PRIOR APPLICATION NUMBER: PCT/EP02/11321
/ PRIOR FILING DATE: 2002-10-09
/ PRIOR APPLICATION NUMBER: EP01124059.5
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 3
/ LENGTH: 2137
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-492-032-3
```

```
Query Match 17.8%; Score 431.4; DB 19; Length 2137;
Best Local Similarity 68.4%; Pred. No. 8.2e-112;
Matches 597; Conservative 0; Mismatches 276; Indels 0; Gaps 0;
```

```
QY 103 CCATGACCATCACTTACAGAACCAAGTGTAAATGCCCGCTTAAAGCTCTTCCGCC 162
Db 210 CCATGACCATCACTTACAGAACCAAGTGTAAATGCCCGCTTAAAGCTCTTCCGCC 269
QY 163 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 222
Db 270 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 329
QY 223 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 282
Db 330 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 389
QY 283 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 342
Db 390 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 449
QY 343 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 402
Db 450 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 509
QY 403 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 462
Db 510 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 569
QY 463 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 522
Db 570 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 629
QY 523 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 582
Db 630 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 689
QY 583 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 642
Db 690 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 749
QY 643 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 702
Db 750 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 809
QY 703 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 762
Db 810 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 869
```

Oy		763	AGTGGGACACTGTATGCTCACTGAAGTTAAGTACCACCGGTATTAACAAGGTGG	822
Dd		870	AATGGGAATGCTCTTTACTATAAGTAAAGCCTCCCTGTATACAGCAAGTGG	929
Oy		823	TGACTGTGGCGGTATACAGCTTCCTTGACTTGTCTACTATTGGGCGGCATTTCTGAAC	882
Dd		930	TGACCATGGCACTGTATACAGCTACTCTGGCTTGCTTATTTGGTGGCGAGTTCCTGAGAC	989
Oy		883	CAGCCAAGGCTTACCTTGCGCATGAGCTGAGCTCGTTGGCCGCTTTACAGTTCCTGC	942
Dd		990	CGGCTCACGGGTACAAAGACACGACCTTAGACCTGTGTGTGCCATCTTCACACCTCTTGC	1049
Oy		943	AGTCTCTTCTATGTGTGGCTGGCTGAAGSTGG	975
Dd		1050	AGTCTCTTCTATGACCGCGCTGGCTCAAGGTAG	1082

RESULT 5
US-10-492-032-30

```

Sequence 30, Application US/10492032
Publication No. US20050049212A1
GENERAL INFORMATION:
APPLICANT: Steuernagel, Arnd
APPLICANT: Bronner, Gunter
APPLICANT: Pritsch, Rudiger
APPLICANT: Buelenberg, Karsten
APPLICANT: Clobesek, Thomas
TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
FILE REFERENCE: 2923-606
CURRENT APPLICATION NUMBER: US/10/492, 032
CURRENT FILING DATE: 2004-04-08
PRIOR APPLICATION NUMBER: PCT/EP02/11321
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
PRIOR FILING DATE: 2001-10-09
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Patentin version 3.2
SEQ ID NO 30
LENGTH: 1530
TYPE: DNA
ORGANISM: Homo sapiens
US-10-492-032-30

```

Query Match 17.7%; Score 431; DB 19; Length 1530;
Best Local Similarity 68.4%; Pred. No. 96-112;
Matches 596; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

Oy		105	ATGACCATCACTTACACAAGCCAGTGGTAAATGCCCGTTAGSCTCTTCTCCGCGCTG	164
Dd		1	ATGACCGGTACCTTACACAGCCCGAGTGGGAAACGCCCGCTTGGTGGCTTCTCCCAAGCTG	60
Oy		165	CTGCTGTCTGTGGCGGGGAGCATCTACAAGCTGCATATAGGAGAATTTCTATCTTCCG	224
Dd		61	CTGCTACTGTGGCGGGAGCATCTACAACATCTGTGGCGAGACTGCTGTCTTCTT	120
Oy		225	CTCTGTCTACTACATATATCCGCTTTTATTATAGSGCTGGCCCTACAGGAACAAGCTG	284
Dd		121	GAGTTCTTACATAGGCGCTAGTGTCTGCCCTAACCGCTTTGTGTGATCCGAAGGCGAAGCC	180
Oy		285	ATGTTTGAAGAACTGACTCTGTATTGGACAAGCTACATACAGCTCATCCCCATTTCTTC	344
Dd		181	TACTTCGAGGAAGCTGTGATTTATTTGTATACAGTATGACAGCTCATCCCTGTCTCTTC	240
Oy		345	GTGCTGGGCTTCTATGTAGCGTGTGTGTAACCGCGTGTGGAAACGATACGAGAAACCTG	404
Dd		241	GTGCTTGGCTTTTATGTAGCGTGTGTGTAACCGCGTGTGGAAACGATACCTATGATCATG	300
Oy		405	CCGTGGCCGACCGGCTCATGAGCTGTGTGTGAGGCTTTGTTCGAAGGCAAGACGAGAA	464
Dd		301	CCGCTGCCGACCGGCTCATGT	360
Oy		465	GAGCGGCTGTGGCGGACGCTATCCGCTACAGCAAACCTGGGCAACGTGCTCATCTCTG	524

[illegible]

```

RESULT 6
US-10-071-766-10/c
: Sequence 10, Application US/10071766
: Publication No. US20020192678A1
: GENERAL INFORMATION:
: APPLICANT: Hua-Ji-Mei Chen
: TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
: FILE REFERENCE: PA-0043 US
: CURRENT APPLICATION NUMBER: US/10/071.766
: CURRENT FILING DATE: 2002-02-07
: NUMBER OF SEQ ID NOS: 144
: SOFTWARE: PERL Program
: SEQ ID NO 10
: LENGTH: 1326
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: incyte ID No. US20020192678A1 1138151.2
US-10-071-766-10

```

	Query Match	14.4%	Score 350.4	DB 13	Length 1326
	Best Local Similarity	99.7%	Pred. No. 8.2e-89		
	Matches 351	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	2053	AGCATTCTTAACTGCTTCTTAATGGGGATGCTTGCAGCAGAGTCTCACCTGTGT	2112		
Db	1326	AGCATTCTTAACTGCTTCTTAATGGGGATGCTTGCAGCAGAGTCTCACCTGTGT	1267		
Qy	2113	GTACACGACGAGCAGCTGATCCAGTCACAGCCATACAGCTGTCCACATGAAAGCGTG	2172		
Db	1266	GTACACGACGAGCAGCTGATCCAGTCACAGCCATACAGCTGTCCACATGAAAGCGTG	1207		
Qy	2173	TCTTCAACAAGCCTGAATCAAATGTTAGCTTAATAGATAAAAATCCACAGCTACTTCAG	2232		
Db	1206	TCTTCAACAAGCCTGAATCAAATGTTAGCTTAATAGATAAAAATCCACAGCTACTTCAG	1147		

QY 2233 CCTTAATGCTTTTATTCATAAAACGTGAAAGTAGACATGAAACCATGGAACATTT 2292
Db 1146 CCTTAATGCTTTTATTCATAAAACGTGAAAGTAGACATGAAACCATGGAACATTT 1087
QY 2293 AACTGACACTCTTGATTCAGAGTCGGGAAACCTTATGTTCTATCTGATATCAAGACAGCA 2352
Db 1086 AACTGACACTCTTGATTCAGAGTCGGGAAACCTTATGTTCTATCTGATATCAAGACAGCA 1027
QY 2353 CACCTTAGATATCTGCGCAACATTAATGAGTTTAAATTAATCAAAATCTCGTT 2404
Db 1026 CACCTTAGATATCTGCGCAACATTAATGAGTTTAAATTAATCAAAATCTCGTT 975

RESULT 7

US-10-492-032-7
Sequence 7, Application US/10492032
Publication No. US20050049212A1
GENERAL INFORMATION:
APPLICANT: Steuernagel, Arnd
APPLICANT: Bronner, Gunter
APPLICANT: Fritsch, Rüdiger
APPLICANT: Eulenbergh, Karsten
TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
FILE REFERENCE: 2923-606
CURRENT FILING DATE: 2004-04-08
PRIOR APPLICATION NUMBER: US/10/492,032
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
PRIOR FILING DATE: 2001-10-09
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.2
SEQ ID NO 7
LENGTH: 1422
TYPE: DNA
ORGANISM: Homo sapiens
US-10-492-032-7

Query Match 14.0%; Score 339.4; DB 19; Length 1422;
Best Local Similarity 64.7%; Pred. No. 1.2e-85;
Matches 505; Conservative 0; Mismatches 276; Indels 0; Gaps 0;

QY 105 ATGACCATCACTTACCAAGCAAGTGGCTAATGCGCTTATGCTCTTCTCCGCTG 164
Db 1 ATGACCGTTTCAATACCTCTCAAGTGGCGAGGCCGCTTGGAGGTTTCTGCGCTG 60
QY 165 CTGCTGCTGCTGGCGGCGAGCATCTACAAAGCTGTATATGCGAGTCTTATCTTCTG 224
Db 61 CTCTCCGCTGAGGAGGAGCATCTACAAAGCTCTCTCAAGGAATCTCTCTTTGGG 120
QY 225 CTCTGCTACTACATCATCGCTTATTTATAGGCTGGCGCTTACAGGAAGCAACAGCTG 284
Db 121 GCTTGTAGCGTGTCTTATGATCACTCAACGCGCTGTGCTGACCCAGGAGCAGAGGTAC 180
QY 285 ATGTTGAGAACTGACTCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTTCTTC 344
Db 181 GTGATGTCTAGTGGCCCGGTACTGCAACCGCTCAAGACCTCATCTCTTGTCTTT 240
QY 345 GTGCTGAGCTTCTAGTGAAGCGTGTGCTGACCCGCTGTGGAAACAGTACGAAACCTG 404
Db 241 GTATTGGTTTCTATGTATCTCTGTGTGAAACCGCTGTGTGTCCTCAAGTACAAAGCATC 300
QY 405 CCGTGGCCCGACCGCTCATGAGCGTGTGTGCGGCTTGTGCAAGGCAAGACAGCA 464
Db 301 CCGCTGCGACAGCTGATGTGTGCTCATCTCGGCTAGGTGACGCGCTGGACAGCGG 360
QY 465 GCGCGGCTGCTGCGGCGGAGCTCATCCGCTTACAGCAACCTGTGGCAACGTTCTCATCTG 524
Db 361 GCGCGCTGCTGCGGCGGAGCTCATCCGCTTACGCAACCTGTGGCGTGTGCTGTGCTG 420
QY 525 CGAGCGTACGACCGGAGTCTCAAGCGCTTCCCGAGGCGCCAGACACTGTGTGACAGCA 584

Db 421 CGCTCGTACAGACCGCGTCTTAAAGCGCTTCCCAACATGAGGACGTGTGAGCA 480
QY 585 GCGTTTATGACTCCGCGAAMACAGAGCATTTGGAGAACTGAGCCTACCAACAATG 644
Db 481 GGTTTCATGTCCCAAGAAAGAGGAAAGAAAGTTTGAAGCCCTGAATCCGACTTCAACAG 540
QY 645 TTCTGGGAGCCCTGGGAGTGTGTTTCCCACTGTCAATGAAGCGTGTGAGGTGCA 704
Db 541 TACTGGGTCCTGCTGTGTGTTTCAACCTGTGGCGGCCAGCCGAGAGGAGCGGCGA 600
QY 705 ATCCGGAACCTATCTGCTCCAGAGCGCTGTGAACAGATGAACAACCTTGCTACAG 764
Db 601 ATACGTGACGATATGCTCTCTGTCTATCTTTGGAAGAGCTGAACAGTACCGAGCGAAG 660
QY 765 TGTGACACCTGTATGCTAGCTAGATGATATGATATCCACTGTGTATACAGGTGTG 824
Db 661 TGACAGATGCTATTCATCACTGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 825 ACTGTGGCGGTATACAGCTTCTTCTGACTGTGTGTGTGTGGGCGAGTTTCTGAACCA 884
Db 721 ACCATAGCCGCTACTCTTTCTTTGCGCTCTCTCTGCTGTGTGTGTGTGTGTGTGTG 780
QY 885 G 885
Db 781 G 781

RESULT 8

US-10-108-260A-1299
Sequence 1299, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. US20040005560A1 full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/10/108,260A
CURRENT FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1299
LENGTH: 2500
TYPE: DNA
ORGANISM: Homo sapiens
US-10-108-260A-1299

Query Match 13.9%; Score 338.2; DB 17; Length 2500;
Best Local Similarity 61.8%; Pred. No. 3.4e-85;
Matches 538; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

QY 105 ATGACCATCACTTACCAAGCAAGTGGCTAATGCGCTTATGCTCTTCTCCGCTG 164
Db 229 ATGACGTGCACTTACTCAAGTAAAGTAGCAATGCAACTTTTGGATTTCATAGTTA 288
QY 165 CTGCTGCTGCTGGCGGCGAGCATCTACAAAGCTGTATATGCGAGTCTTATCTTCTG 224
Db 289 CTCTCAAGTGGAGGAGCAATCTACAAAGCTGTACAGGAATTTATGTTTGTCT 348
QY 225 CTCTGCTACTACATCAACCGCTTATTTATAGGCTGGCGCTTACGGAAGCAACAGCTG 284
Db 349 GTTCTTTATACAGCAATTAATGTTGTGTGACAGATTCTTATCAAGAGTCCAAAAAGT 408
QY 285 ATGTTGAGAACTGACTGTATTTGCGACAGCTACATCCAGCTCATCCCATTTCTTC 344
Db 409 TACTTTGAAAAATTATCAATTTATCTGTGACATATGCTGAACAAATTCAGTAACCTTT 468
QY 345 GTGCTGAGCTTCTAGTGAAGCGTGTGCTGACCCGCTGTGTGAAACAGTACAGAACTG 404
Db 469 GTGCTTGGGTTTATATGTTACTCTGTGTGTAACCATGATGTGGAACCAAGTTTGTGAA 528
QY 405 CCGTGGCCCGACCGCTCATGAGCGTGTGTGCGGCTTGTGCAAGGCAAGACGAGCA 464
Db 529 CCTGGCCAGAGAGGTATATGTTCTCTCATCTTACAGATGTTACGGAAGGAGCAGAGCAC 588

Qy	465	GGCGGCTGCTCGGGCGCAGCTATCTCCGTAACGGCAACTGAGGCAAGTGTATCTCG	524
Db	589	GGGGCGCTGCTTTAAGAAAGACGCTGATGCGCTACGTCATATCTCACTCCCTGCTATCTTT	648
Qy	525	CGACGCGTCAGCACCGCCAGCTCATCAAGCGCTTCCCGACGGCCGACCACTGGTGCACGA	584
Db	649	CGCTCGGTAGACACTGCTGTGTACAAAAGATTTCCCAATGAGACCAAGTGGTGAACGA	708
Qy	585	GGCTTTATGACTCCGGGCAAGAACACAGCATTTGGAGAACTGAGCTTACACACAATG	644
Db	709	GGTTTATGACAACAGATGAAAGAAATTAATTCACCACTCAAGTCTCTCATCTGAAA	768
Qy	645	TTTCGGGCGCCCTGGGTGTGTTGGCCAACTGTAAATGAAGCGTGGCTGTGAGAGTGA	704
Db	769	TATTTGGTTCATTTCATCTGTGGTTGGAAATCTTGGAACTTAAGCCCGGAAATGAAGTAGA	828
Qy	705	ATCCGGGAGCCCTATCTGCTCCAGAGCCGCTGAACGAGATGAACACTTGGGCTACTAG	764
Db	829	ATCAGAGCAAGTGTGTGATCTGCATCATTTGATGACTGAAATGAATCATACCGCTTTGG	888
Qy	765	TGTGACACCTGTATGCCCTACGACTGTGATTAGTATCCCATGCTGTATATACAGGTGTG	824
Db	889	TGCAGCCTTATTTCGTTATGATGACCTGGGTTGGGATTCGCTGTTTAAACCCAGGTGTC	948
Qy	825	ACTGTGCGGTGTACAGCTTCTTCTCTGACTTGTCTAATTGTGGCGGCACTTTCTGAACCA	884
Db	949	ACTGTGCTGTATATACCTTCTTCTTTCGAGTGCATTTGACGCGCAATTTTGGATCCC	1008
Qy	885	GCCAAAGGCTTACCTTGGCCATGAGTGTGAACCTCGTTGTGCCGCTTTACAGTTCTCTGAG	944
Db	1009	ACCAAAAGGCTACGCGAGGCAATGACTTGGATCTTTTACATTTCCCATCTTCAACCTCCTCAAA	1068
Qy	945	TTCTTCTTCTAATGTGAGCTGGGCTGAAGGAG	975
Db	1069	TTCTTCTTCTAATGAGATGGCTTAAAGTAG	1099

```

RESULT 9
US-10-492-032-5
; Sequence 5, Application US/10492032
; Publication No. US20050049212A1
; GENERAL INFORMATION:
; APPLICANT: Steuernagel, Arnd
; APPLICANT: Bronner, Gunter
; APPLICANT: Pritsch, Rudiger
; APPLICANT: Eulenberg, Karsten
; APPLICANT: Closssek, Thomas
; TITLE OF INVENTION: Bacteriophin and Bacteriophin Homologous Proteins Involved in the
; TITLE OF INVENTION: Regulation of Energy Homeostasis
; FILE REFERENCE: 2923-606
; CURRENT APPLICATION NUMBER: US/10/492, 032
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: PCT/EP02/11321
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: EP01124059.5
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 5
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-492-032-5

Query Match      13.9%; Score 336.8; DB 19; Length 2028;
Best Local Similarity 61.8%; Pred. No. 7,7e-85;
Matches 536; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

07      105  ATGACATCATCTTACACAAAGCAAGTGGCTATGCCCCGGTTAGGCTTCCTTCCGCCCTG 164
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       1   ATGATGTCATCTACTCTCCAGTAAGTAGCAAAATGCAACTTTTGGATTTCAATAGGTTA 60

```

QY	165	CTGCTGTCCTGAGCGGGGAGCATCTACAAAGCTGTATATGCGAGTTCTTAACTTCCTG	224
Db	61	CTCTCTCAAGTGGAGGCGACGACTACACAACTACTGTACAGGGAAATTTATGTTTTGCT	1220
QY	225	CTCTGCTACTACATCATCCGCTTTTATTTATATAGCTGCGCCCTCA	CGGAGAAACAAGCTG 284
Db	121	GTTCTTTTATACAGCAATAAAGTTTGGTATATACAGATGTTTACTTACAGAGTCCAAAAAGT	180
QY	285	ATGTTTGGAAACTGACTCTGTATTTGCGACAGCTACATCCAGTCAATCCCCATTTTCTTC	344
Db	181	TACTTTGAAAAATTAATCAATTTACTGTGACAAATATGCTGAAACAAATTTCCAGTAACCTTT	240
QY	345	GTGCTGGAGCTTCTACGTGACGCTGTGTCGTAACCCCTGTGTGAAACCATGACGAAACTG	404
Db	241	GTGCTGTGGGTTTATATGTACTCTGTGTATGTGAACCGATGCTGAAACCAATTTTGGAAATTTG	300
QY	405	CCGTGGCCCGACCCGCTCATGAGCGCTGTGTGCGGCTTCTGTGAAAGCAAGACGACAA	464
Db	301	CCCTGGCCAGACAGGCTTAATGTTCTCTACTCTACAGTGTTCACGGAAGCCGACGAGCAC	360
QY	465	GGCGGCTGTGTGCGGCGGACGCTCAATCCGCTACCGCAACCTGGGCAACGTCGCTCAATCG	524
Db	361	GGGCGGCTGTGAAAGACGCTGATGGCTACGTCATCTCACTCCCTGCTCAATCTT	420
QY	525	CGACAGCTCAGACACCGCAGTCTACAAAGCGCTTCCCCAGCGCCAGCAGCACTGTGTGACAA	584
Db	421	CGCTCGGTGACACTGCTGTGTACAAAAGATTTTCCACAAATGGACCAAGCTGTGTGAAGCA	480
QY	585	GAGCTTATGACTCCGGCAGAACACACAGCAGTTGGAGAACTGAGCCTACCAACAACATG	644
Db	481	GGTTTATAGACAAACAGATGAAAGAAATTAATTAACACACTCAAGTCTCTCATCTGAAA	540
QY	645	TTCTGGGTGGCCCTGGGTGTGTTGCCAACTGTACATGAAGCGTGGCTTGGAGGTCGA	704
Db	541	TATTTGGTTCATTTCAATCATCTGTTTGGAAATCTTGCACTTAAAGCCCGGAATGAAGGTAGA	600
QY	705	ATCCGGGACCCATATCTGCTCCAGAGCGCTGTGAAAGCATGAACAACCTTGTGCTACTCAG	764
Db	601	ATCAGAGACAGTGTGATCTGTCAATCATTTGATGACGAAATGATATGATATACGCTCTTGG	660
QY	765	TGTGGACACTGTATGCTCTACGACTGTGATTAGTATCCACTGTGTATACACAGGTGTG	824
Db	661	TGCAGCCCTTAATTCGGTTATGACCTGGGTGGGAATTCCTCGTGTTCACCCAGGTGTGC	720
QY	825	ACTGTGGGGGTATACAGCTTCTCTCTGACTGTCTTATAGTTGGGGCGAGATTTTCGAACCA	884
Db	721	ACTCTGTGTCTATACCTTCTCTTGTGTGCTCTGATTTGAAACGCCAGATTTTGGATCCC	780
QY	885	GCCAAAGGCTTACCCCTGGACATGAGCTGTGACCTGTGTGTGCCGTCTTCAAGTTCTGTGAG	944
Db	781	ACCAAAAGGCTACGCAAGGGACATGACTTTGGATCTTTTACATTTCCCATCTTCAACCCCTCTACAA	840
QY	945	TTCTTTCTTCTATATGTTGGCTGGCTGAAG	972
Db	841	TTCTTTCTTCTATATGAGATGCGCTTAAG	868

RESULT 10
US-09-814-353-19504/C
; Sequence 19504, Application US/09814353
; Publication No. US2003016583A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21

```

; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19504
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19504

```

```

Query Match      8.8%; Score 213; DB 10; Length 620;
Best Local Similarity 66.4%; Pred. No. 8.6e-50;
Matches 306; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

```

```

QY 103 CCATGACCATCATTAACACAGCCCAAGTGGCTAATGCGCTTAGGCTCTCCGCC 162
DB 464 CCATGACGGTTTCATACACTCTCAAGGCGGAGCCCGCTTCGAGAGTTCTCGGCC 405
QY 163 TGCCTGCTGCTGGGGGACACATCTACAGCTGCTATAGCGAGTTCTTAATCTCC 222
DB 404 TCCCTCTCGCTGAGGGGAGACATCTACAGCTCTCTACAGAGAAATTCCTCCCTTGG 345
QY 223 TGCCTGCTACTACATCACTCCGCTTATTTATTAAGGCTGAGCCCTCAAGAGCAAGC 282
DB 344 GGGCGCTTGAAGCTGTGCTAGCACTACCTACCGGCTGCTGTAACCGAGAGAGGT 285
QY 283 TGATGTTGAGAACTGACTGTATAGCGACAGCTACATCCAGCTCAATCCCAATTCCT 342
DB 284 AGGTATAGCTCAGAGTGGCCCGGTACTGCAACCGCTCAGAGACCTCATTCCTTGTCT 225
QY 343 TCGCTGAGGCTTCAAGCGTGAAGCTGCTGTAAGCCCGCTGGTGAACCAATGACGAACC 402
DB 224 TTGATTTGGGTTTCTATGTAAGCTCTGTAAGACCGGTGATGTCCTCCAGTACCAAGCA 165
QY 403 TCCGCTGAGCCGACCGCTCAAGAGCTGTGCTGAGGCTTGTGCAAGGCAAGAGCAGC 462
DB 164 TCCGCTGACAGACAGCTGATGAGCGTCACTGCGCTAGAGTGCACCGCTGAGACAGC 105
QY 463 AAGCGCGCTGCTGCGGAGCAGCTCATCGCTACGCCCACTGGGCAACGCTGCTCATCC 522
DB 104 GGGGCGGCTGCTGCGGAGCAGCTCATCGCTACGCCCACTGGGCTTGGTGTGCTGC 45
QY 523 TGCGCAGGCTCAGACCGCATCTTACAGAGGCTTCCCGCAG 563
DB 44 TGGCCTCGGTACAGACCGGCTGCTTAAGCGCTTCCCGAC 4

```

```

RESULT 11
US-10-198-846-11070
; Sequence 11070, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846

```

```

; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11070
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1287, 1288, 1289, 1290, 1291, 1292
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

```

```

Query Match      7.8%; Score 189.2; DB 14; Length 1292;
Best Local Similarity 62.4%; Pred. No. 7.7e-43;
Matches 296; Conservative 0; Mismatches 178; Indels 0; Gaps 0;

```

```

QY 349 TGGGCTTCTACGTAAGCTGTGCTGTAACCCGCTGCTGGAACCAATGACAGAACCTGCCGT 408
DB 318 TAGGTTTATGTAAGTACTGTGTAGTAACCGATGCTGGAACCAATTTGTAATTTCCCT 377
QY 409 GGGCCGACCGCTCATAGAGCTGTGTGCGGCTTTCGTAAGGCAAGCAAGCAAGGCC 468
DB 378 GGCAGACAGGCTTAATGTTCTCATCTCTAGCAGTGTTCAGAGAGCAAGCAAGGCC 437
QY 469 GCGTCTGCGCGGCGACGCTCATCCGCTACGCAACCTGGAGCAAGCTCTATCTCGCA 528
DB 438 GCGTCTTAAAGAGCGCTGATGCGCTACGCTCATCTCCTCGCTCATCTTTCGCT 497
QY 529 GCGTACGACCGGAGTCTTACAGCGCTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 588
DB 498 CGGTGAGCACTGTGTGTATCAAAAGATTTCACAAATGACAGAGGTTGAAGCAGGT 557
QY 589 TTATGACTCGGCAAGCAACAGCACTGGAAGAACTGAGGCTTACCAACAATGTTCT 648
DB 558 TTATGACAAAGATGAAGAAATTTATTCACACACTCAAGTCTCTCATGAAATATT 617
QY 649 GGGTGGCGGTGCTGTGTTGCAACCTGCTCAATGAGGCTGCTGAGAGTCCATCC 708
DB 618 GGGTTCATTCATCTGTGTGTAAGTCTTGCAACTTAAACCCGGAATGAAGGTAGATCA 677
QY 709 GGGACCTTATCCGCTCCAGAGCTGCTGAGAGATGAACCTTGCCTACTCAGTGTG 768
DB 678 GAGACAGTGTATGATGCAATCATTTGATGACTGAATGATGATGATGATGATGATGATG 737
QY 769 GACACTGTATGCTTACAGCTGATTAATGATCCACTGCTGTATACACAGGTG 822
DB 738 GCCTTATTCGTTATGACTGGGTGGGATTCGCTGCTTACACCCAGGTAG 791

```

```

RESULT 12
US-09-880-107-2174/C
; Sequence 2174, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2174
; LENGTH: 1198

```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L20941
US-09-860-107-2174
```

```
Query Match          7.2%; Score 175; DB 9; Length 1198;
Best Local Similarity 100.0%; Pred. No. 8.3e-39;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2230 CAGCCTTTATGCTTTTATTCATTAATAAAGCTGTAAGCTGAGTGAACATTGGAAACA 2289
DB 1198 CAGCCTTTATGCTTTTATTCATTAATAAAGCTGTAAGCTGAGTGAACATTGGAAACA 1139
QY 2290 TTTAAGCTAGAGCTGAGTTCAGAGTCCGGAACCTTGTCTCTGTAATCCAGAGACAG 2349
DB 1138 TTTAAGCTAGAGCTGAGTTCAGAGTCCGGAACCTTGTCTCTGTAATCCAGAGACAG 1079
QY 2350 CCAGACCTTATGATCTGCTGAGTCCCAACTAATGATTAATAATACAAATACTCGTT 2404
DB 1078 CCAGACCTTATGATCTGCTGAGTCCCAACTAATGATTAATAATACAAATACTCGTT 1024
```

```
RESULT 13
US-10-027-632-134530/c
; Sequence 134530, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134530
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134530
```

```
Query Match          6.6%; Score 161.4; DB 13; Length 578;
Best Local Similarity 99.4%; Pred. No. 4.3e-35;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 582 GCAGCCTTTATGATCTGCGGAGAACACAGAGCTTGAGAAAGAGCTTACACACAC 641
DB 183 GCAGCCTTTATGATCTGCGGAGAACACAGAGCTTGAGAAAGAGCTTACACACAC 124
QY 642 ATGTTCTGGGTCCTGCGGTGTGTTGCCAACCTGTCAATGAAGGCGTGGCTTGAAGT 701
DB 123 ATGTTCTGGGTCCTGCGGTGTGTTGCCAACCTGTCAATGAAGGCGTGGCTTGAAGT 64
QY 702 CGAATCCGGGACCTTATCTGCTCCAGAGCTTGTCTGAACGAGA 744
DB 63 CGAATCCGGGACCTTATCTGCTCCAGAGCTTGTCTGAACGAGA 21
```

```
RESULT 14
US-10-027-632-134529/c
; Sequence 134529, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134530
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134530
```

```
Query Match          6.6%; Score 161.4; DB 17; Length 578;
Best Local Similarity 99.4%; Pred. No. 4.3e-35;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 582 GCAGCCTTTATGATCTGCGGAGAACACAGAGCTTGAGAAAGAGCTTACACACAC 641
DB 183 GCAGCCTTTATGATCTGCGGAGAACACAGAGCTTGAGAAAGAGCTTACACACAC 124
QY 642 ATGTTCTGGGTCCTGCGGTGTGTTGCCAACCTGTCAATGAAGGCGTGGCTTGAAGT 701
DB 123 ATGTTCTGGGTCCTGCGGTGTGTTGCCAACCTGTCAATGAAGGCGTGGCTTGAAGT 64
QY 702 CGAATCCGGGACCTTATCTGCTCCAGAGCTTGTCTGAACGAGA 744
DB 63 CGAATCCGGGACCTTATCTGCTCCAGAGCTTGTCTGAACGAGA 21
```

```
RESULT 15
US-10-027-632-134529/c
; Sequence 134529, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
```

; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 134529
 ; LENGTH: 748
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-134529

Query Match 6.2%; Score 149.4; DB 13; Length 748;
 Best Local Similarity 90.9%; Pred. No. 1,3e-31;
 Matches 170; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 558 CCCAGCGCCCGACACCTGTGCAAGCAGGCTTATGACTCCGCGAGACACAAGCAGTTG 617
 |||||
 Db 712 CCCACCGCTTTTTCACCTCCACTCTGCGAGGCTTATGACTCCGCGAGACACAAGCAG-TCG 654
 QY 618 GAGAACTGAGCTTACCAACAACATGTTCTGGGTGCCCTGGGTGTGTTGCCAACCTG 677
 |||||
 Db 653 GAGAACTGAGCTTACCAACAACATGTTCTGGGTGCCCTGGGTGTGTTGCCAACCTG 594
 QY 678 TCAATGAAGCGCTGTGGCTTGGAGGTGGAATCCGCGACCTTATCTGCTCCAGAGCTGCTG 737
 |||||
 Db 593 TCAATGAAGCGCTGTGGCTTGGAGGTGGAATCCGCGACCTTATCTGCTCCAGAGCTGCTG 534
 QY 738 AACGAGA 744
 |||||
 Db 533 AACGTGA 527

Search completed: March 26, 2005, 22:07:42
 Job time : 1362.06 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 22:47:44 ; Search time 387.221 seconds
(without alignment)
10264.201 Million cell updates/sec

Title: US-09-622-964a-4
Perfect score: 2429
Sequence: 1 cagggagctccaccagccta.....aaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1954.8	80.5	2171	4	US-09-949-016-4901
2	642.2	26.4	7609	4	US-09-949-016-16644
3	642.2	26.4	18075	4	US-09-949-016-16643
4	358.4	14.8	1462	4	US-09-949-016-4902
5	188.4	7.8	601	4	US-09-949-016-173056
6	161.4	6.6	601	4	US-09-949-016-173051
7	161.4	6.6	601	4	US-09-949-016-173052
8	128	5.3	601	4	US-09-949-016-173045
9	89	3.7	49416	4	US-09-949-016-123932
10	89	3.7	49416	4	US-09-949-016-15234
11	88.6	3.6	601	4	US-09-949-016-46055
12	88.6	3.6	43086	4	US-09-949-016-133055
13	88.4	3.6	678533	4	US-09-949-016-14577
14	88.4	3.6	10827	4	US-09-949-016-14578
15	87.8	3.6	10827	4	US-09-949-016-12297
16	87.8	3.6	11939	4	US-09-949-016-13232
17	87.8	3.6	17605	4	US-09-949-016-13599
18	87.8	3.6	25231	4	US-09-949-016-12041
19	87.8	3.6	25431	4	US-09-949-016-13234
20	87.8	3.6	50836	4	US-09-949-016-16722
21	87.4	3.6	27663	4	US-09-949-016-16160
22	87.4	3.6	28109	4	US-09-949-016-14329
23	87.4	3.6	31573	4	US-09-949-016-15628
24	87.4	3.6	31573	4	US-09-949-016-16327
25	87.4	3.6	35629	4	US-09-949-016-15786
26	87.2	3.6	275110	4	US-09-949-016-12706
27	87.2	3.6	275110	4	US-09-949-016-16070

28	86.6	3.6	601	4	US-09-949-016-121853	Sequence 121853, A
29	86.4	3.6	9608	4	US-09-949-016-16706	Sequence 16706, A
30	86.2	3.5	601	4	US-09-949-016-28769	Sequence 28769, A
31	86.2	3.5	601	4	US-09-949-016-28770	Sequence 28770, A
32	86.2	3.5	601	4	US-09-949-016-52283	Sequence 52283, A
33	86.2	3.5	601	4	US-09-949-016-52285	Sequence 52285, A
34	86.2	3.5	601	4	US-09-949-016-52286	Sequence 52286, A
35	86.2	3.5	601	4	US-09-949-016-52287	Sequence 52287, A
36	86.2	3.5	601	4	US-09-949-016-77763	Sequence 77763, A
37	86.2	3.5	601	4	US-09-949-016-77764	Sequence 77764, A
38	86.2	3.5	601	4	US-09-949-016-117618	Sequence 117618, A
39	86.2	3.5	601	4	US-09-949-016-126371	Sequence 126371, A
40	86.2	3.5	601	4	US-09-949-016-126372	Sequence 126372, A
41	86.2	3.5	601	4	US-09-949-016-126373	Sequence 126373, A
42	86.2	3.5	601	4	US-09-949-016-168504	Sequence 168504, A
43	86.2	3.5	601	4	US-09-949-016-168552	Sequence 168552, A
44	86.2	3.5	601	4	US-09-949-016-168600	Sequence 168600, A
45	86.2	3.5	16236	4	US-09-949-016-16236	Sequence 16236, A

ALIGNMENTS

```
RESULT 1
US-09-949-016-4901
; Sequence 4901, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR APPLICATION NUMBER: 2000-04-14, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4901
; LENGTH: 2171
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4901

Query Match      80.5%; Score 1954.8; DB 4; Length 2171;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 2169; Conservative 0; Mismatches 2; Indels 203; Gaps 1;

QY      17  CTTAGTGCAGACCTTCTGTGGATCATCGGACCCAGTGAACCCCACTGACCCAG 76
DB      1  CCTAGTGCAGACCTTCTGTGGATCATCGGACCCAGTGAACCCCACTGACCCAG 60

QY      77  CCCACCTGTCGACGCCACCTGTCGACCATCATCATCTTACCAAGCCAAAGGCGTAA 136
DB      61  CCCACCTGTCGACGCCACCTGTCGACCATCATCATCTTACCAAGCCAAAGGCGTAA 120

QY      137 TGCCCGCTTAGGAGCTTCTTCCCGCTGTCGTCGTCGGGGGAGACATCTCAAGCT 196
DB      121 TGCCCGCTTAGGAGCTTCTTCCCGCTGTCGTCGTCGTCGGGGGAGACATCTCAAGCT 180

QY      197 GCTATATGCGAGTCTTAACTCTTCTGCTCTGCTACTATCATCATCCGCTTATTTATAG 256
DB      181 GCTATATGCGAGTCTTAACTCTTCTGCTCTGCTACTATCATCATCCGCTTATTTATAG 240

QY      257 GCTGGCCCTTACCGGAAACAAACAGCTGATTTGAAACTGACTCTGTAATTCGACAG 316
DB      241 GCTGGCCCTTACCGGAAACAAACAGCTGATTTGAAACTGACTCTGTAATTCGACAG 300

QY      317 CTACATCAGCTCATCCCATTTCTCTGTCGGGCTTCTACCTGACGCTGTCGAC 376
```

Db 301 CTACATCCAGCTCATCCCATTTCTTCTGCTGGCTCTTACGAGAGCTGGTGGTAC 360
Qy 377 CCGCTGTGGAAACCAATACGAGAACCTGCGGTGGCCGACCGCTCATGAGCTGTGTGTC 436
Db 361 CCGCTGTGGAAACCAATACGAGAACCTGCGGTGGCCGACCGCTCATGAGCTGTGTGTC 420
Qy 437 GGGCTTCTGAGAGCAAGAGACGAGCAAGGCGGTGGTGGGCGACAGCTCATCCGCTA 496
Db 421 GGGCTTCTGAGAGCAAGAGACGAGCAAGGCGGTGGTGGGCGACAGCTCATCCGCTA 480
Qy 497 CGGCAACTGGGCAACGTGCTCATCTGCGAGCGTGAACAACGAGCTTACAAGCGCTT 556
Db 481 CGGCAACTGGGCAACGTGCTCATCTGCGAGCGTGAACAACGAGCTTACAAGCGCTT 540
Qy 557 CCCCAGCCCGCAGCACTGTGTGAAGAGGCTTTATGATCCGCGAACAACAAGCAAT 616
Db 541 CCCCAGCCCGCAGCACTGTGTGAAGAGGCTTTATGATCCGCGAACAACAAGCAAT 600
Qy 617 GAGAAATGAGCTTACCAACAACATGTTCTGGGTCCCTGGGTGGTGGTGGTGGTGGT 676
Db 601 GAGAAATGAGCTTACCAACAACATGTTCTGGGTCCCTGGGTGGTGGTGGTGGTGGT 660
Qy 677 GTCAATGAAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 736
Db 661 GTCAATGAAGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 720
Qy 737 GAACGAGATGAACCTTGGCTACTCAGTGTGAACAACCTGTATGCTTACGATGATG 796
Db 721 GAACGAGATGAACCTTGGCTACTCAGTGTGAACAACCTGTATGCTTACGATGATG 780
Qy 797 TATCCAGTGTGTATACAGAGTGTGATCTGTGGGTGGTGGTGGTGGTGGTGGTGGT 856
Db 781 TATCCAGTGTGTATACAGAGTGTGATCTGTGGGTGGTGGTGGTGGTGGTGGTGGT 840
Qy 857 TCTAGTGGGCGGCACTTCTGAACCAAGCAAGGCTTACCTGGCGAATGAGCTGACCT 916
Db 841 TCTAGTGGGCGGCACTTCTGAACCAAGCAAGGCTTACCTGGCGAATGAGCTGACCT 900
Qy 917 CGTTGGGCGGCTTCTGACCTTCTGCAATCTTCTTATGTTGGCTGGTGAAGTGGG 976
Db 901 CGTTGGGCGGCTTCTGACCTTCTGCAATCTTCTTATGTTGGCTGGTGAAGTGGG 954
Qy 977 CCTCTCAAGGCGCTGCTGGGCTGAGAGCATGGCGAAGGGTCAATGGCCAGAGCTGCT 1036
Db 955 ----- 954
Qy 1037 TGAAGCAGAGATGCAAGTGTCAAGAAAGAGGTCTCAAGGTGAAGAAAGCAGCGGTG 1096
Db 955 ----- 954
Qy 1097 GTGGCGCAACCTGTATATCCAGCTACTCGGGAAGGCTGAGGAGAGATCGCTTGAACC 1156
Db 955 ----- 954
Qy 1157 CGGGAGGCGAGGTTGTGTGGCAGAGAGCTCATCAACCTTTGGAGAGATGATGAT 1216
Db 955 -----GTTGGCAGAGAGCTCATCAACCTTTGGAGAGATGATGAT 997
Qy 1217 GATTTTGAACCACTGGAATTTGTGAAGAAATTTGAGAGTTCCTGTTGGCTGTGGAT 1276
Db 998 GATTTTGAACCACTGGAATTTGTGAAGAAATTTGAGAGTTCCTGTTGGCTGTGGAT 1057
Qy 1277 GAGATGCAACAGGACTGCTCGATGAGACCGGACATGTACTGGAATTAAGCCCGACCA 1336
Db 1058 GAGATGCAACAGGACTGCTCGATGAGACCGGACATGTACTGGAATTAAGCCCGACCA 1117
Qy 1337 CAGCCCCCTTACAGAGTGTCTCGCCCACTTCCGTGAGCTCTTTATGGGCTTCAAC 1396
Db 1118 CAGCCCCCTTACAGAGTGTCTCGCCCACTTCCGTGAGCTCTTTATGGGCTTCAAC 1177
Qy 1397 TTCAATCATGAGCTGAACAAAGAGATGAGTTCAGCCCAATGAGAGACAGAGAG 1456
Db 1177 TTCAATCATGAGCTGAACAAAGAGATGAGTTCAGCCCAATGAGAGACAGAGAG 1456

Db 1178 TTCAATCATGAGCTGAACAAAGAGATGAGTTCAGCCCAATGAGAGACAGAGAG 1237
Qy 1457 GATGCTCAAGCTGTGATATGAGCGCTTCTTGAAGCTTGAAGTCAATGATCAATGCT 1516
Db 1238 GATGCTCAAGCTGTGATATGAGCGCTTCTTGAAGCTTGAAGTCAATGATCAATGCT 1297
Qy 1517 CCGAGGCAAACTCAAGGACAAACTATGAGGCGCAAGAGGAAATCCCTTCTCAAGAG 1576
Db 1298 CCGAGGCAAACTCAAGGACAAACTATGAGGCGCAAGAGGAAATCCCTTCTCAAGAG 1357
Qy 1577 GGCCTGCGCAAAACCAAGAGCGCAAAACAGAACCTTGAAGGCGCAGAGACCAAG 1636
Db 1358 GGCCTGCGCAAAACCAAGAGCGCAAAACAGAACCTTGAAGGCGCAGAGACCAAG 1417
Qy 1637 GCTTGAAGCTTGAAGGCGTGAAGCGCTTCAAGTGTGAGGCGCAATGATCAAGAGCG 1696
Db 1418 GCTTGAAGCTTGAAGGCGTGAAGCGCTTCAAGTGTGAGGCGCAATGATCAAGAGCG 1477
Qy 1697 TACTACAGTGGCCCAAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTGAGACCA 1756
Db 1478 TACTACAGTGGCCCAAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTGAGACCA 1537
Qy 1757 TCAAGCGCGTCAAAAGCTTCAAGTGTCAAGGCTATGACACCAAGACAAAGCTTAAAG 1816
Db 1538 TCAAGCGCGTCAAAAGCTTCAAGTGTCAAGGCTATGACACCAAGACAAAGCTTAAAG 1597
Qy 1817 ACTGTAGTCTGGGCGCAAGAAAGTTTGAATGCTCTGAGAGAGGATGGGCGCTTG 1876
Db 1598 ACTGTAGTCTGGGCGCAAGAAAGTTTGAATGCTCTGAGAGAGGATGGGCGCTTG 1657
Qy 1877 ATGAGACCCCAAAATATCTCAAGTGAAGAGAAACCTGTGAGATTTAACTGAACGAT 1936
Db 1658 ATGAGACCCCAAAATATCTCAAGTGAAGAGAAACCTGTGAGATTTAACTGAACGAT 1717
Qy 1937 ATGCAAGATCCCGAAATCACTCAAGAAACCTTTGGAACATCAACCAACCACTA 1996
Db 1718 ATGCAAGATCCCGAAATCACTCAAGAAACCTTTGGAACATCAACCAACCACTA 1777
Qy 1997 CACACTACACTCAAGATCAATGATCTTATGAGGCTTGAAGAAACAGGATGAGCA 2056
Db 1778 CACACTACACTCAAGATCAATGATCTTATGAGGCTTGAAGAAACAGGATGAGCA 1837
Qy 2057 CATTCCTAACCTGCTTCCAAATGAGGAGTCTTGGCAGGCGAGTCTCACTGTGTGAC 2116
Db 1838 CATTCCTAACCTGCTTCCAAATGAGGAGTCTTGGCAGGCGAGTCTCACTGTGTGAC 1897
Qy 2117 ACCAGCAGGACATGATCAAGCTCAAGCATTACAGTGTCCACATGAAGACGTGCT 2176
Db 1898 ACCAGCAGGACATGATCAAGCTCAAGCATTACAGTGTCCACATGAAGACGTGCT 1957
Qy 2177 ACAACAGCTGAATCAAAATGTTAGCTTATAGTAAATCCCAAGCTTACCTGAGCTT 2236
Db 1958 ACAACAGCTGAATCAAAATGTTAGCTTATAGTAAATCCCAAGCTTACCTGAGCTT 2017
Qy 2237 TAAAGCTTTTATTCATPAAAACTGTAAAGCTTGAAGCTTGAAGCAATTTAAGT 2296
Db 2018 TAAAGCTTTTATTCATPAAAACTGTAAAGCTTGAAGCTTGAAGCAATTTAAGT 2077
Qy 2297 CAGACTCGAATTAAGTGGGAAACCTTATGTTCTATCTGAATCCAGACAGCCACCC 2356
Db 2078 CAGACTCGAATTAAGTGGGAAACCTTATGTTCTATCTGAATCCAGACAGCCACCC 2137
Qy 2357 TTAGTATATGCTCCCAATCAATGATGTTAATAAA 2390
Db 2138 TTAGTATATGCTCCCAATCAATGATGTTAATAAA 2171

RESULT 2
US-09-949-016-16644/c
; Sequence 16644, Application us/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: C1001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 16644
;; LENGTH: 7609
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-16644

Query Match 26.4%; Score 642.2; DB 4; Length 7609;
Best Local Similarity 99.5%; Pred. No. 4e-157;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1405 CAGCCTGAACAAAGAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGATGCTCA 1464
DB 7382 CAGCCTGAACAAAGAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGATGCTCA 7323
QY 1465 CGCTGGCATCTTGGCGCGCTTCTTAAGGCTGCAAGTCCCATATCACCATTCTCCAGGGC 1524
DB 7322 CGCTGGCATCTTGGCGCGCTTCTTAAGGCTGCAAGTCCCATATCACCATTCTCCAGGGC 7263
QY 1525 AAACCTCAAGAGCAAACTACTGTGGCCCAAGAGGGAATCCCTTCTCAAGAGGCTGCTCC 1584
DB 7262 AAACCTCAAGAGCAAACTACTGTGGCCCAAGAGGGAATCCCTTCTCAAGAGGCTGCTCC 7203
QY 1585 CAAAAACCAAGAGCAGCCCAACAGAAAGTTAGGGGCGAGAGAGCAACAAAGGCTGGAA 1644
DB 7202 CAAAAACCAAGAGCAGCCCAACAGAAAGTTAGGGGCGAGAGAGCAACAAAGGCTGGAA 7143
QY 1645 GCTTAAGGCTGTGAGACGCTTCAAGTCTGCGCCCACTGATTCAGAGGCTGACTACTACAG 1704
DB 7142 GCTTAAGGCTGTGAGACGCTTCAAGTCTGCGCCCACTGATTCAGAGGCTGACTACTACAG 7083
QY 1705 TGCCCCACAGACGCGCCCTCAGCCCACTCCCATGTTCTTCCCTTAGAAGCATCAGAGGCC 1764
DB 7082 TGCCCCACAGACGCGCCCTCAGCCCACTCCCATGTTCTTCCCTTAGAAGCATCAGAGGCC 7023
QY 1765 GTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAGAGCAAAAGCTTAAAGACTGTGAG 1824
DB 7022 GTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAGAGCAAAAGCTTAAAGACTGTGAG 6963
QY 1825 TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCGCTTGATGAGACA 1884
DB 6962 TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCGCTTGATGAGACA 6903
QY 1885 CCCAGAAGTATCTCAAGTGAAGAGGAAACCTGTGAGTTTAACCTGACGATATGCGCAGA 1944
DB 6902 CCCAGAAGTATCTCAAGTGAAGAGGAAACCTGTGAGTTTAACCTGACGATATGCGCAGA 6843
QY 1945 GATCCCCGAAATATCACTCAAGAAACCTTTGGAACAATACCAACCAACATATACACACTAC 2004
DB 6842 GATCCCCGAAATATCACTCAAGAAACCTTTGGAACAATACCAACCAACATATACACACTAC 6783
QY 2005 ACTCAAGATCAATGATCTTATTTGGGCTTTGGAAGACGAGATG 2051
DB 6782 ACTCAAGATCAATGATCTTATTTGGGCTTTGGAAGACGAGATG 6736

RESULT 3
US-09-949-016-16643
; Sequence 16643, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

;; APPLICANT: VENTER, J. Craig et al.
;; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;; FILE REFERENCE: C1001307
;; CURRENT APPLICATION NUMBER: US/09/949,016
;; CURRENT FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/241,755
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 16643
;; LENGTH: 18075
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-16643

Query Match 26.4%; Score 642.2; DB 4; Length 18075;
Best Local Similarity 99.5%; Pred. No. 6.6e-157;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1405 CAGCCTGAACAAAGAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGATGCTCA 1464
DB 13879 CAGCCTGAACAAAGAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGATGCTCA 13938
QY 1465 CGCTGGCATCTTGGCGCGCTTCTTAAGGCTGCAAGTCCCATATCACCATTCTCCAGGGC 1524
DB 13939 CGCTGGCATCTTGGCGCGCTTCTTAAGGCTGCAAGTCCCATATCACCATTCTCCAGGGC 13998
QY 1525 AAACCTCAAGAGCAAACTACTGTGGCCCAAGAGGGAATCCCTTCTCAAGAGGCTGCTCC 1584
DB 13999 AAACCTCAAGAGCAAACTACTGTGGCCCAAGAGGGAATCCCTTCTCAAGAGGCTGCTCC 14058
QY 1585 CAAAAACCAAGAGCAGCCCAACAGAAAGTTAGGGGCGAGAGAGCAACAAAGGCTGGAA 1644
DB 14059 CAAAAACCAAGAGCAGCCCAACAGAAAGTTAGGGGCGAGAGAGCAACAAAGGCTGGAA 14118
QY 1645 GCTTAAGGCTGTGAGACGCTTCAAGTCTGCGCCCACTGATTCAGAGGCTGACTACTACAG 1704
DB 14119 GCTTAAGGCTGTGAGACGCTTCAAGTCTGCGCCCACTGATTCAGAGGCTGACTACTACAG 14178
QY 1705 TGCCCCACAGACGCGCCCTCAGCCCACTCCCATGTTCTTCCCTTAGAAGCATCAGAGGCC 1764
DB 14179 TGCCCCACAGACGCGCCCTCAGCCCACTCCCATGTTCTTCCCTTAGAAGCATCAGAGGCC 14238
QY 1765 GTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAGAGCAAAAGCTTAAAGACTGTGAG 1824
DB 14239 GTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAGAGCAAAAGCTTAAAGACTGTGAG 14298
QY 1825 TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCGCTTGATGAGACA 1884
DB 14299 TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCGCTTGATGAGACA 14358
QY 1885 CCCAGAAGTATCTCAAGTGAAGAGGAAACCTGTGAGTTTAACCTGACGATATGCGCAGA 1944
DB 14359 CCCAGAAGTATCTCAAGTGAAGAGGAAACCTGTGAGTTTAACCTGACGATATGCGCAGA 14418
QY 1945 GATCCCCGAAATATCACTCAAGAAACCTTTGGAACAATACCAACCAACATATACACACTAC 2004
DB 14419 GATCCCCGAAATATCACTCAAGAAACCTTTGGAACAATACCAACCAACATATACACACTAC 14478
QY 2005 ACTCAAGATCAATGATCTTATTTGGGCTTTGGAAGACGAGATG 2051
DB 14479 ACTCAAGATCAATGATCTTATTTGGGCTTTGGAAGACGAGATG 14525

RESULT 4
US-09-949-016-4902/c
; Sequence 4902, Application US/09949016
; Patent No. 6812339

```
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 4902
/ LENGTH: 1462
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-4902
```

```
Query Match      14.8%; Score 358.4; DB 4; Length 1462;
Best Local Similarity 99.7%; Pred. No. 2.2e-83;
Matches 359; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 2045 AGGATGAGCAGCATTTCTTAACCTGCTTCTAATGAGGATGCTTGGCAGCCAGGTCCTC 2104
DB 1385 AGGATGAGCAGCATTTCTTAACCTGCTTCTAATGAGGATGCTTGGCAGCCAGGTCCTC 1326
OY 2105 ACCTGTGTACACAGCAGGAGCAGTATCCAGTCAGCAGCAGCAGTATCCAGTCACTGA 2164
DB 1325 ACCTGTGTACACAGCAGGAGCAGTATCCAGTCAGCAGCAGCAGTATCCAGTCACTGA 1266
OY 2165 AGAAGCTGTCTTCAACAAGCCTGAATCAATGTTAGTTAATAGATAAATATCCAGAC 2224
DB 1265 AGAAGCTGTCTTCAACAAGCCTGAATCAATGTTAGTTAATAGATAAATATCCAGAC 1206
OY 2225 TACTTCAGCCTTAAATGCTTTATTCATTAATGTTAATGTTAATGTTAATGTTAATGTTA 2284
DB 1205 TACTTCAGCCTTAAATGCTTTATTCATTAATGTTAATGTTAATGTTAATGTTAATGTTA 1146
OY 2285 AAACATTAACTGAGACTGATTCAGAGTGGGAGACCTAGTCTTCAATGTAATGTA 2344
DB 1145 AAACATTAACTGAGACTGATTCAGAGTGGGAGACCTAGTCTTCAATGTAATGTA 1086
OY 2345 GACAGCCACACCTTAGTATATGCCCCAATCAATGATTAAATCAATACTCGTT 2404
DB 1085 GACAGCCACACCTTAGTATATGCCCCAATCAATGATTAAATCAATACTCGTT 1026
```

```
RESULT 5
US-09-949-016-173056
/ Sequence 173056, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 173056
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
```

```
US-09-949-016-173056
```

```
Query Match      7.8%; Score 188.4; DB 4; Length 601;
Best Local Similarity 99.5%; Pred. No. 3.3e-39;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 2045 AGGATGAGCAGCATTTCTTAACCTGCTTCTAATGAGGATGCTTGGCAGCCAGGTCCTC 2104
DB 412 AGGATGAGCAGCATTTCTTAACCTGCTTCTAATGAGGATGCTTGGCAGCCAGGTCCTC 471
OY 2105 ACCTGTGTACACAGCAGGAGCAGTATCCAGTCAGCAGCAGCAGTATCCAGTCACTGA 2164
DB 472 ACCTGTGTACACAGCAGGAGCAGTATCCAGTCAGCAGCAGCAGTATCCAGTCACTGA 531
OY 2165 AGAAGCTGTCTTCAACAAGCCTGAATCAATGTTAGTTAATAGATAAATATCCAGAC 2224
DB 532 AGAAGCTGTCTTCAACAAGCCTGAATCAATGTTAGTTAATAGATAAATATCCAGAC 591
OY 2225 TACTTCAGCC 2234
DB 592 TACTTCAGCC 601
```

```
RESULT 6
US-09-949-016-173051
/ Sequence 173051, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 173051
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-173051
```

```
Query Match      6.6%; Score 161.4; DB 4; Length 601;
Best Local Similarity 99.4%; Pred. No. 3.8e-32;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 582 GCAGGCTTTATGACTCCGAGCAACAAGAGTTGAGAACTGAGCTTACACACAC 641
DB 321 GCAGGCTTTATGACTCCGAGCAACAAGAGTTGAGAACTGAGCTTACACACAC 380
OY 642 ATGTTCTGAGTCCGCGGAGTGTGTTGCAACCTGTCAATGAAGCGTGTGAGGT 701
DB 381 ATGTTCTGAGTCCGCGGAGTGTGTTGCAACCTGTCAATGAAGCGTGTGAGGT 440
OY 702 CGAATCCGGAGCCCTATCTGTCTCCAGAGCCTGTGAACGAGA 744
DB 441 CGAATCCGGAGCCCTATCTGTCTCCAGAGCCTGTGAACGAGA 483
```

```
RESULT 7
US-09-949-016-173052
/ Sequence 173052, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 173052
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-173052

Query Match 6.6%; Score 161.4; DB 4; Length 601;
Best Local Similarity 99.4%; Pred. No. 3.8e-32;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 582 GCAGCCTTATGACATCCGGGAGAACAAAGAGCTTGAGAACTGAGCTTACACACAC 641
Db 30 GCAGCCTTATGACATCCGGGAGAACAAAGAGCTTGAGAACTGAGCTTACACACAC 89
Oy 642 ATGTTCTGGGTCCTGCGGTGTGTTTGCACACTGTCAATGAGGCGTGGAGAGT 701
Db 90 ATGTTCTGGGTCCTGCGGTGTGTTTGCACACTGTCAATGAGGCGTGGAGAGT 149
Oy 702 CGAATCCGGAGACCTATCTCTGCCAGAGCTGTGTAACGAGA 744
Db 150 CGAATCCGGAGACCTATCTCTGCCAGAGCTGTGTAACGAGA 192

RESULT 8
US-09-949-016-173045
Sequence 173045, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 173045
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-173045

Query Match 5.3%; Score 128; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 2e-23;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 69 GACCCAAACCCACCTGCTGACAGCCCACTGCGGCAATGACATCTATACCAAGCCAA 128
Db 474 GACCCAAACCCACCTGCTGACAGCCCACTGCGGCAATGACATCTATACCAAGCCAA 533
Oy 129 GTGGCTAATGCCCGGCTTGTGCTTCCCGCCTGTGCTGTGTGTGGCGGGGAGCATC 188
Db 534 GTGGCTAATGCCCGGCTTGTGCTTCCCGCCTGTGCTGTGTGTGGCGGGGAGCATC 593
Oy 189 TACAAGCT 196
Db 189 TACAAGCT 196

Db 594 TACAAGCT 601

RESULT 9
US-09-949-016-123932
Sequence 123932, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 123932
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-123932

Query Match 3.7%; Score 89; DB 4; Length 601;
Best Local Similarity 86.7%; Pred. No. 3.1e-13;
Matches 98; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 1081 AAGAGCCAGGCGGTGGGCGGACACCTGTAATCCAGCTACTCGGAGGCTGAGGCG 1140
Db 422 AATTAAGCCAGGCGGTGGGCGGACACCTGTAATCCAGCTACTCGGAGGCTGAGGCG 481
Oy 1141 GAGATCGCTTGAACCCGGAGGCGGAGGTTGTGTGCGACAGAGCTCATCA 1193
Db 482 TAGATCGCTTGAACCCGGAGGCGGAGGTTGTGTGCGACAGAGCTCATCA 534

RESULT 10
US-09-949-016-15234/C
Sequence 15234, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO: 15234
LENGTH: 49416
TYPE: DNA
ORGANISM: Human
US-09-949-016-15234

Query Match 3.7%; Score 89; DB 4; Length 49416;
Best Local Similarity 86.7%; Pred. No. 3.8e-12;
Matches 98; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Oy 1081 AAGAGCCAGGCGGTGGGCGGACACCTGTAATCCAGCTACTCGGAGGCTGAGGCGAG 1140
Db 17799 AATTAAGCCAGGCGGTGGGCGGACACCTGTAATCCAGCTACTCGGAGGCTGAGGCGAG 17740

Qy 1141 GAGAATCGCTTGAACCCGGGAGCGCGAGGTTGTGGTGGCAGAGCAAGCTCATCA 1193
|||||
Db 17739 TAGAATCGCTTGAACCCGGGAGCGCGAGGTTGTGGTGGCAGCAAGATCACACCA 17687

RESULT 11
US-09-949-016-46055/C
; Sequence 46055, Application US/09949016

```

: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001107
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 46055
:
: LENGTH: 601
:
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-46055

```

	Query Match	Similarity	Score	DB	Length
Best Local	91.3%	Pred. No.	3.9e-13		
Matches	94	Conservative	0	Mismatches	9
				Indels	0
				Gaps	0
Qy	1081	AAAGAGCGACGCGCGTGTGGCGCAACA	CGTGAATCCGACGCTACGCGGAGCGTGAAGGAG	1141	
Db	583	AAATTAGCGGCGCGTGTGGCGCAACGCTGTATCCGACTGATGTGTGGGGGCGTGAAGGAG	524		
Qy	1141	GAGCAATCGCTTGAATCCCGGAGAGCGCGAGGTTTGTGTGGCGAG	1183		
Db	523	GAGCAATCGCTTGAATCCCGGAGAGCGCGAGGTTTGTGTGGCGAG	481		

RESULT 12
 US-09-949-016-13059/C
 ; Sequence 13059, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: C1001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13059
 ; LENGTH: 43086
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-13059

Query Match	3.6%	Score 88.6;	DB 4;	Length 43086;
Best Local Similarity	91.3%;	Pred. No. 4.5e-12;		
Matches 94;	Conservative 0;	Mismatches 9;	Indels 0;	Gaps 0;

Qy	Db	Qy	Db
108	108	1141	10185
AAAGCAGCCAGGGGTGTGTGTGCGC	AAATTGACCGGGGGTGTGTGTGCGC	GAGAAATCGCTTGAATCCCGGAGCGG	GAGAAATCGCTTGAATCCCGGAGCGG
CAACCTGTAATCCAGCTATCTCGGAGG	CAACCTGTAATCCAGCTATCTCGGAGG	AGGTTGTGTGTGCGCAAG	AGGTTGTGTGTGCGCAAG
GTAGGACAG	GTAGGACAG		
10245	10245		
AAATTGACCGGGGGTGTGTGTGCGC	AAATTGACCGGGGGTGTGTGTGCGC		
CAACCTGTAATCCAGCTATCTCGGAGG	CAACCTGTAATCCAGCTATCTCGGAGG		
GTAGGACAG	GTAGGACAG		
10185	10185		
GAGAAATCGCTTGAATCCCGGAGCGG	GAGAAATCGCTTGAATCCCGGAGCGG		
AGGTTGTGTGTGCGCAAG	AGGTTGTGTGTGCGCAAG		
10143	10143		

RESULT 13
US-09-949-016-14577

```

: Sequence 14577, Application US/09949016
: Patent No. 6812319
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ. ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14577
: LENGTH: 678533
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(678533)
: OTHER INFORMATION: n = A,T,C or G
: IS-09-949-016-14577

```

	Query Match	Best Local Similarity	Matches	Conservative	Score	DB 4	Pred.	No. 2.5e-11	Mismatches	Indels	Gaps
Qy	1079	AGAAACACCACAGGGGTGTTGGCGACACTGTATCCCACTTCTCGGAGGCTTGAGGC	1138		3.6%						
Db	499918	AAAAATTACCCACAGCGGTGTGGCGCACCTGTATTCCTCAGATGTGTAGAC	499977		73.4%				41	0	0
Qy	1139	AGGAGATCGCTTGAACCCGGAGGCGAGGTTGTGTGSCAAGCAAGCTTCATCAACCCC	1198								
Db	499978	AGGAGATCGCTTGAACCCGGAGGCAAGGTTGTGTGAGCCAAAGTATCAATCAGGCC	500037								
Qy	1199	TTTGGAGAGGATGATGATGATTTTGAACAACCT	1232								
Db	500038	ATTGCACTCCAGCTTGGGCAACMAAGGAAGAAT	500071								

```

RESULT 14
US-09-949-016-14578
: Sequence 14578, Application US/09949016
: Patent No. 6812338
GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012

```


This Page Blank (uspro)

Db	1	GTGCNAAGCCAGTACTATCACTACAGAAAGTACGCAATCCCGCTTCGTTGCTT	60
Qy	61	CTCGCCCTCCCTCCGTGTCGCGAGGAGAGATCTACAGAGTCTGTATGAGAAATTCCT	120
Db	61	CTCGTCCTCCCTCCGTGTCGCGAGGAGAGATCTACAGAGTCTGTATGAGAAATTCCT	120
Qy	121	TGTCTTCATATTCCTTACTATATTCATCCGTCGACTCAAGAAATGTTCTCTCGAGTGA	180
Db	121	TGTCTTCATATTCCTTACTATATTCATCCGTCGACTCAAGAAATGTTCTCTCGAGTGA	180
Qy	181	TCACAGAGCTGTTGTTTGAAGAGCTGTCGTACTGCGACAGCTCAATTCAGCTCATCC	240
Db	181	TCACAGAGCTGTTGTTTGAAGAGCTGTCGTACTGCGACAGCTCAATTCAGCTCATCC	240
Qy	241	TATATCCCTGCTTCTGGGTTTCTATGTTTACATGGTGTGAGCCGCTGTGTGAGCCAGTA	300
Db	241	TATATCCCTGCTTCTGGGTTTCTATGTTTACATGGTGTGAGCCGCTGTGTGAGCCAGTA	300
Qy	301	CGAGAACCTTGCCTGCGCCGACCGCTCATGATCCAGTGTCTAGCTTCTGTGAGGCGAA	360
Db	301	CGAGAACCTTGCCTGCGCCGACCGCTCATGATCCAGTGTCTAGCTTCTGTGAGGCGAA	360
Qy	361	GGATGAGAGAGGCGGTTTGTGGGGGCGACGCTCATCCGTCAGCCATCTGTGGCCAACT	420
Db	361	GGATGAGAGAGGCGGTTTGTGGGGGCGACGCTCATCCGTCAGCCATCTGTGGCCAACT	420
Qy	421	GCTCATCTGCGAGCATGACGACTCGGTCTACAGCGCTTCCACCTTCAACACCT	480
Db	421	GCTCATCTGCGAGCATGACGACTCGGTCTACAGCGCTTCCACCTTCAACACCT	480
Qy	481	GGGTCTAGAGAGGTTTATACCCATGAGGAAATTAAGAGTTCAGAAATGGGCTTACC	540
Db	481	GGGTCTAGAGAGGTTTATACCCATGAGGAAATTAAGAGTTCAGAAATGGGCTTACC	540
Qy	541	ACAACAACATTTCTGGGTCCTCGGTGTGTTGTGCAATGTGTAAGAGGCGCTTACT	600
Db	541	ACAACAACATTTCTGGGTCCTCGGTGTGTTGTGCAATGTGTAAGAGGCGCTTACT	600
Qy	601	TGGAGGTCAATTCGGGACACCGCTCTGTCAGAGCTGTAAGTGAAGTGTGTACTTT	660
Db	601	TGGAGGTCAATTCGGGACACCGCTCTGTCAGAGCTGTAAGTGAAGTGTGTACTTT	660
Qy	661	GGTACTAGTGTGACAGTGTATGCTTACAGACTGTAAGTGAAGTGTGTACTTT	720
Db	661	GGTACTAGTGTGACAGTGTATGCTTACAGACTGTAAGTGAAGTGTGTACTTT	720
Qy	721	ACAGGTGTGACAGTGTGAGTATACAGCTTTTCTTGCATGTGTGATCCGAGGCGAGTT	780
Db	721	ACAGGTGTGACAGTGTGAGTATACAGCTTTTCTTGCATGTGTGATCCGAGGCGAGTT	780
Qy	781	TCTGAACCCAAACAAAGGACTACCCAGGCGATGAGATGTGTTGTGCTTCTTAC	840
Db	781	TCTGAACCCAAACAAAGGACTACCCAGGCGATGAGATGTGTTGTGCTTCTTAC	840
Qy	841	AATCTGTCAATTTCTATCTACATGAGGCTGCTGAGGTGAGAAAGAGCTCATCAACC	900
Db	841	AATCTGTCAATTTCTATCTACATGAGGCTGCTGAGGTGAGAAAGAGCTCATCAACC	900
Qy	901	CTTGGGAGAGACATGATGATTTTGAAGTAACTGATCATTTGACAGAAACCTGTGAGT	960
Db	901	CTTGGGAGAGACATGATGATTTTGAAGTAACTGATCATTTGACAGAAACCTGTGAGT	960
Qy	961	GTCCTGTGTCCTGTGATGAGTACAGAACTTGCCTCCATGGAACGTGACATGTA	1020
Db	961	GTCCTGTGTCCTGTGATGAGTACAGAACTTGCCTCCATGGAACGTGACATGTA	1020
Qy	1021	CTGGAAGAGAGAGCGCTTACGCGCCCTTACAGCTGTCTTCTGCAAGGTCTGCGCGCA	1080
Db	1021	CTGGAAGAGAGAGCGCTTACGCGCCCTTACAGCTGTCTTCTGCAAGGTCTGCGCGCA	1080
Qy	1081	TTCCCTCATGGGCTCCACCTTCAACATCAAGCTTAAAGAAAGACTTGAAGCTTTGCTC	1140

Db	1081	TTCCCTCATGGGCTCCACCTTCAACATCAAGCTTAAAGAAAGACTTGAAGCTTTGCTC	1140
Qy	1141	AAAAGAGAGGCTGACAGATTAAGAAAGAGGAGCTATATAGGACGACATAGGCTCTT	1200
Db	1141	AAAAGAGAGGCTGACAGATTAAGAAAGAGGAGCTATATAGGACGACATAGGCTCTT	1200
Qy	1201	CTTAGAGCTGCAACCCAAACAACTACATCTTCCCTTGAAGAGACTTAAAGCAAACTATT	1260
Db	1201	CTTAGAGCTGCAACCCAAACAACTACATCTTCCCTTGAAGAGACTTAAAGCAAACTATT	1260
Qy	1261	GTTGTTCTAAGAACCCCTCCCTCCGAAAGGCGAGTGTAAAGATGCCAACGAAAAACAGAA	1320
Db	1261	GTTGTTCTAAGAACCCCTCCCTCCGAAAGGCGAGTGTAAAGATGCCAACGAAAAACAGAA	1320
Qy	1321	AGATGTCTGAAATTTTAAAGGCTGTGACTTCTTGAATGTGTTCCAAGTTTAAAGGAG	1380
Db	1321	AGATGTCTGAAATTTTAAAGGCTGTGACTTCTTGAATGTGTTCCAAGTTTAAAGGAG	1380
Qy	1381	AGGCTCCCATTTGTGGCCCAAGGCAAGCAGCAGCCCTTACTGAGCAGTCAAGCCTTC	1440
Db	1381	AGGCTCCCATTTGTGGCCCAAGGCAAGCAGCAGCCCTTACTGAGCAGTCAAGCCTTC	1440
Qy	1441	CAGTTCAAGACAGAGTATGAGGCTTCCACAGATTACCAAGAAATCTGTCAATGAAAA	1500
Db	1441	CAGTTCAAGACAGAGTATGAGGCTTCCACAGATTACCAAGAAATCTGTCAATGAAAA	1500
Qy	1501	GAAACCTGTGAGTTTAACTTGAACATTCAGAGAGCCCAAGAAATCTTCAACAGCG	1560
Db	1501	GAAACCTGTGAGTTTAACTTGAACATTCAGAGAGCCCAAGAAATCTTCAACAGCG	1560
Qy	1561	CCGTTTGGACACAGATGTCAACCAATTAAGAGCTTAAATGAAGAGATGACAGTCTTA	1620
Db	1561	CCGTTTGGACACAGATGTCAACCAATTAAGAGCTTAAATGAAGAGATGACAGTCTTA	1620
Qy	1621	TCCCTACAGGATGAAGCTGTGACCAACCTGTTCTATGATGATGCTTCAACAGCTTG	1680
Db	1621	TCCCTACAGGATGAAGCTGTGACCAACCTGTTCTATGATGATGCTTCAACAGCTTG	1680
Qy	1681	GCCCTGACCTTGAAGATATGCCAGAGGAGCACTGACCCAGTCAAAAGGACACAAAGCAG	1740
Db	1681	GCCCTGACCTTGAAGATATGCCAGAGGAGCACTGACCCAGTCAAAAGGACACAAAGCAG	1740
Qy	1741	ACACCCAGAGAGTGTTCACAGACAGTCTAGCATGTAACTCAGAACCAAGATCTTAA	1800
Db	1741	ACACCCAGAGAGTGTTCACAGACAGTCTAGCATGTAACTCAGAACCAAGATCTTAA	1800
Qy	1801	TAGTCTGTGCTGAAAAACCTGTATTTTAAAGATCTTCCCAACTAAGAGATTTAATAA	1860
Db	1801	TAGTCTGTGCTGAAAAACCTGTATTTTAAAGATCTTCCCAACTAAGAGATTTAATAA	1860
Qy	1861	CGTGAATATTTCTTTAAGGTGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1916
Db	1861	CGTGAATATTTCTTTAAGGTGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA	1916

RESULT 2
AY450427 1904 bp mRNA linear ROD 24-JUN-2004
LOCUS AY450427
DEFINITION Mus musculus vitelliniform macular dystrophy 2 (Vmd2) mRNA, complete cds.
ACCESSION AY450427
VERSION AY450427.1 GI:41612988
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1904)
Stohr, H., Marguier, A., Nanda, I., Schmid, M. and Weber, B. H.
TITLE Three novel human VMD2-like genes are members of the evolutionary highly conserved RPP-TM family
JOURNAL Eur. J. Hum. Genet. 10 (4), 281-284 (2002)
MEDLINE 22027749

PUBMED 12032738
REFERENCE 2 (bases 1 to 1904)
AUTHORS Kraemer, F., Stoeck, H. and Weber, B.H.F.
TITLE Cloning and characterization of the murine Vmd2 Ref-1m gene family
JOURNAL Cytogenet. Genome Res. 105 (1), 107-114 (2004)
AUTHORS 3 (bases 1 to 1904)
AUTHORS Weber, B.H.F. and Kraemer, F.
TITLE Direct Submision
JOURNAL Submitted (27-OCT-2003) Human Genetics, University of Wuerzburg, Am
Hubland, Biocentre, Wuerzburg 97074, Germany
FEATURES
Source
1..1904
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/chromosome="19"
/tissue_type="testis"
1..1904
/gene="Vmd2"
37..1692
/gene="Vmd2"
/codon_start=1
/product="Vitellogenin macular dystrophy 2"
/protein_id="AA09922.1"
/db_xref="GI:41612989"
/translation="MTITVTNKKVANNARLGSFSSLLCMRGSITKYLIVGEPLVFIPLY
SIRLRYNVLSSDOQLPEKALYCDSTIQIPISPVIGFVTVTVSRRMSQYENLPM
PDRIMVQSVFEGKDEBGRLLRLRLIRYALIGQVLLIRSTISVYKRPFLNHLVLA
GPMTHGKHQKQLKGLPNTFPMVWVPMANLMSKAYLGGRLRDVLLQSLMNEVCTLR
TQCCGLAYDWISILPLYTVVTVAVVSFLACILGKQFLPNKDYGVGHMLVVPF
TILQFLPYMGWLKVAEOLINFGEDDDFEFTNAILDRNLOVSLISVGHOMDLPMER
DMVNEAPAPOPPTAASARSBRHSFMGSTFVILSKKDLIELMSKEDADTKKESGYSS
TIGFELGLQPKNTYRLPLKDLTKILCSNPLLEBQCCDANKDKWPKFGDLPLK
VPRKRGSHCGPAPSPHSPTGSDPSDGTGDSPTDYGCKMKKTVFENLIPES
SPTEHLQORRLDQSTNLIQALMKHAESYPYRDAEGTKPYLYE"

ORIGIN
Query Match 97.8%; Score 1874.8; DB 10; Length 1904;
Base Local Similarity 99.9%; Pred. No. 0;
Matches 1876; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Query 1 GTGCCAAGCCATGATCATCTACACAAACAAAGTACCAATGCCCGCTCGGTTCTT 60
Db 27 GTGCCAAGCCATGATCATCTACACAAACAAAGTACCAATGCCCGCTCGGTTCTT 86
Query 61 CTGCTCCCTCTCTCTGCTGCTGCGAGAGCAGCATCTACAAGCTGCTGTAAGAAATTCCT 120
Db 87 CTGCTCCCTCTCTCTGCTGCTGCGAGAGCAGCATCTACAAGCTGCTGTAAGAAATTCCT 146
Query 121 TGTCTTCAATTTCTCTACTATTCATTCCTGCTGAGACTCTACAGAAATGTTCTTCTGAGTGA 180
Db 147 TGTCTTCAATTTCTCTACTATTCATTCCTGCTGAGACTCTACAGAAATGTTCTTCTGAGTGA 206
Query 181 TCACAGAGCTGTTGTTGAGAAGCTGCTGTAAGTGCAGACAGTACATTCAGCTCATCC 240
Db 207 TCACAGAGCTGTTGTTGAGAAGCTGCTGTAAGTGCAGACAGTACATTCAGCTCATCC 266
Query 241 TATAATCTTCTGTTCTGAGTTTCTATGTTACATTTGCTGAGTGAAGCGCTGGTGAAGCAGTA 300
Db 267 TATAATCTTCTGTTCTGAGTTTCTATGTTACATTTGCTGAGTGAAGCGCTGGTGAAGCAGTA 326
Query 301 CGAATACTTGGCGTGGCCGACCGCTCATATGATCCAGTGTCTAGCTTCTGAGAGGGAA 360
Db 327 CGAATACTTGGCGTGGCCGACCGCTCATATGATCCAGTGTCTAGCTTCTGAGAGGGAA 386
Query 361 GGATGAGGAGAGCGCTTGTGCGGCGCAGCGCTCATCCGCTACAGCAATCTGCGCAAGT 420
Db 387 GGATGAGGAGAGCGCTTGTGCGGCGCAGCGCTCATCCGCTACAGCAATCTGCGCAAGT 446
Query 421 GCTCATCTGCGAGAGCATCAGCACTCGGTCTACAAAGCGCTTTCCTTCAACACACT 480
Db 447 GCTCATCTGCGAGAGCATCAGCACTCGGTCTACAAAGCGCTTTCCTTCAACACACT 506

Query 481 GGTGCTAGCAGGTTTATGACCCATGCGGGAACATTAAGCAGTTTGAGAAATTGGGCTTACC 540
Db 507 GGTGCTAGCAGGTTTATGACCCATGCGGGAACATTAAGCAGTTTGAGAAATTGGGCTTACC 566
Query 541 ACACAACATCTTGGGCTGCTGGGTGTGTTGTTGCACTTGTCAATGAAGGCTTATCT 600
Db 567 ACACAACATCTTGGGCTGCTGGGTGTGTTGTTGCACTTGTCAATGAAGGCTTATCT 626
Query 601 TGGAGTGGAAATCCGGGACACCGCTGCTCAGAGCTGATGAATGAGGTGTACTT 660
Db 627 TGGAGTGGAAATCCGGGACACCGCTGCTCAGAGCTGATGAATGAGGTGTACTT 686
Query 661 GCGTACTCAGTGTGAGACAGCTGTATGCTTACAGCTGATTAAGTATCCCATTTGTTACAC 720
Db 687 GCGTACTCAGTGTGAGACAGCTGTATGCTTACAGCTGATTAAGTATCCCATTTGTTACAC 746
Query 721 ACAGTGGTGAACAGTGCAGTATACAGCTTTTCTTGCATGCTTGAATCGGAGGCAATT 780
Db 747 ACAGTGGTGAACAGTGCAGTATACAGCTTTTCTTGCATGCTTGAATCGGAGGCAATT 806
Query 781 TCTGAACCCAAACAAAGACATACCCAGGCAATGAGATGAGTGTGCTGTCTTAC 840
Db 807 TCTGAACCCAAACAAAGACATACCCAGGCAATGAGATGAGTGTGCTGTCTTAC 866
Query 841 AATCCTCAATTTCTTATTTCTACATGAGCTGGCTGAAGGTGAGCAACAGCTCATCAACC 900
Db 867 AATCCTCAATTTCTTATTTCTACATGAGCTGGCTGAAGGTGAGCAACAGCTCATCAACC 926
Query 901 CTTGCGGAGAGACGATGATGATTTTGAAGATTAACATGATCTGAACAGAACTTGCAGCT 960
Db 927 CTTGCGGAGAGACGATGATGATTTTGAAGATTAACATGATCTGAACAGAACTTGCAGCT 986
Query 961 GTCCCTGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 987 GTCCCTGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1046
Query 1021 CTGGAACGAGCAGCGCTCTAGCGGCTTACACAGCTGCTTCTGCAAGTCTCCGCGCA 1080
Db 1047 CTGGAACGAGCAGCGCTCTAGCGGCTTACACAGCTGCTTCTGCAAGTCTCCGCGCA 1106
Query 1081 TTCCTTATGAGGCTTCAACCTTCAACATCAAGCTTAAGAAAGACATTAAGGCTTGTGTC 1140
Db 1107 TTCCTTATGAGGCTTCAACCTTCAACATCAAGCTTAAGAAAGACATTAAGGCTTGTGTC 1166
Query 1141 AAAAGAGAGGCTTACACGATTAAGAAAGAGAGTGTATGACGACCAATAGGCTGTT 1200
Db 1167 AAAAGAGAGGCTTACACGATTAAGAAAGAGAGTGTATGACGACCAATAGGCTGTT 1226
Query 1201 CTTAGACCTGCAACCCAAACATCAACATCTTCCCTTGAAGACATTAAGACCAACATATT 1260
Db 1227 CTTAGACCTGCAACCCAAACATCAACATCTTCCCTTGAAGACATTAAGACCAACATATT 1286
Query 1261 GTGTTCTTAAGAACCCCTCTCTGAAAGGCAAGTGTAAAGTGCACCAAGAAAAACAGAA 1320
Db 1287 GTGTTCTTAAGAACCCCTCTCTGAAAGGCAAGTGTAAAGTGCACCAAGAAAAACAGAA 1346
Query 1321 AGATGCTGGAATTTAAGGCTGTGACCTTGTGAATGTGTTCGAAGGTTTAAGAGAG 1380
Db 1347 AGATGCTGGAATTTAAGGCTGTGACCTTGTGAATGTGTTCGAAGGTTTAAGAGAG 1406
Query 1381 AGGCTCCCATTTGAGGCCCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 1440
Db 1407 AGGCTCCCATTTGAGGCCCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGG 1466
Query 1441 CAGTTCAAGACACAGTGTAGGCTTCTCAAGATTACCAAGAAATCTGTCAATGAATAA 1500
Db 1467 CAGTTCAAGACACAGTGTAGGCTTCTCAAGATTACCAAGAAATCTGTCAATGAATAA 1526
Query 1501 GAAACCTGTGAGATTTAATTTGAACATTCAGAGAGCCCAAGAAATCTTCAACAGCG 1560
Db 1527 GAAACCTGTGAGATTTAATTTGAACATTCAGAGAGCCCAAGAAATCTTCAACAGCG 1586


```

Db      563 GCTCATCTCGGAGCATCAGACACTCGCTACAGGCGTTCCAGTCTTCAACACCT 622
Qy      481 GGTCTTAGCAGATTTTATGACCCATGGGAACTATAGAGTTGACAGAGTGGGCTTACC 540
Db      623 GATCAAGCAGGCTTTATGACTCTCGGGGAATATAGAGTTGACAGAGTGGGCTTACC 682
Qy      541 ACACAACATCTTCTGGGTCCTGGGTGGTTGGTTGGCCAACTGTTCATGAAGAGCTTATCT 600
Db      683 ACACAACATCTTCTGGGTGACCTTGGGTGGTTGGTTGGCCAACTGTTCATGAAGAGCTTATCT 742
Qy      601 TGAAGTGAATCCGGGACACCGTCTGCTCAGAGCCCTGATGAATGAATGATGATCTTT 660
Db      743 TGAAGTGAATCAGGAGCACTGTCTGCTCAGAGCCCTGATGAATGAATGATGATGATCTTT 802
Qy      661 GCGTACTCAGTGTGACAGCTGTATGCTTACGATCTGATTAAGTATCCCTTGGTGTACAC 720
Db      803 GCGTACTCAGTGTGACACCTATATGCTTATGACTGTATCAATATCCCATTTGGTGTACAC 862
Qy      721 ACAGGTGATGACAGTGGCAGTATACAGCTTTTCTTGCATGCTGATCGGAGGAGT 780
Db      863 ACAGGTGATGACAGTGGGAGTTTACAGCTTTTCTTGCATGCTGATCGGAGGAGT 922
Qy      781 TCTGAACCCAAACAGAGCTACCCAGGCCATGAGATGATGATGATGATGATGATGATGATGAT 840
Db      923 TCTGAACCCACTCAAGAGCTACCCAGGCCATGAGATGATGATGATGATGATGATGATGATGAT 982
Qy      841 AATCTGCAATTTCTTATTTCTACATGGGCTGCTGATGATGATGATGATGATGATGATGATGAT 900
Db      983 AATCTGCAATTTCTTATTTCTACATGGGCTGCTGATGATGATGATGATGATGATGATGATGAT 1042
Qy      901 CTTCGGGAGGAGCAGTATGATTTTGAAGCTATACGATCATCTGATCAAGAACTGCAAGT 960
Db      1043 CTTCGGGAGGAGTATGATTTTGAAGCTATGATCATCTGATCATCTGATCAAGAACTGCAAGT 1102
Qy      961 GTCCCTGTGTCCGTGATGGATGACACAGAACTTGCCTCCATGGAACCTGACATGTA 1020
Db      1103 GTCCCTGTGTCTGTGATGGATGACACAGAGTGTCTCCCTGGAGGAGACCTGTATA 1162
Qy      1021 CTGGAACGAGGAGCGCTCAGCGCCCTTACACAGCTGCTTCTGTCAGGCTTCCCGGCA 1080
Db      1163 CTGGAATGAGGCGCGCTCAGCGCCCTTACACAGCTGCTTCTGTCAGGCTTCCCGGCA 1222
Qy      1081 TTCTTCATGGGCTCCACTTCATCATAGGCTTAAGAAAGAACTAGAGCTTTGGTC 1140
Db      1223 ATCTTCATGGGCTCCACTTCATCATAGGCTTAAGAAAGAAAGCTAGAG----- 1274
Qy      1141 AAAAGAGAGGCTGACACGATAAGAAAGAGTGTATAGCAGACCATAGGCTGCTT 1200
Db      1275 -GAGGAGAGTACACTTATGACAAAGAAAGAAATGATTAAGGACAGCATTTGACGCTT 1333
Qy      1201 CTTCGACCTGCAACCCCAAAATCATACATCTTCCCTGAAAGACTTTAAAGACCAATATT 1260
Db      1334 CTTCGACCTGCAATCTCATGACTACATCTTCCCAAGCACAATTAAGACCAATCTGTT 1393
Qy      1261 GTGTCTTAAGA---ACCCCTCTCGAAGGCGAGTGAAGATGCAACGAAAGAAACCA 1317
Db      1394 GTGTCTTAAGAAGACCCCTTCTCTGAGACAGCTTAAGATATACAGCAGAAAGACCA 1453
Qy      1318 GAAAGATGTCTGAAATTTAAAGGCTTGAAGCTTCTTGAATATGTCTCAAGGTTTAAAG 1377
Db      1454 GAAAGATG---ATGTCTGAAACCAAGAGACCTCTTCAGACCTATACCAAGTTTAAAG 1510
Qy      1378 GAGAGGCTCCATTTGTGGCCACAGGACCCAGACCA-----CCCTAC 1422
Db      1511 GCGCAAGTTACCGCAGGCGCCACAGCCACTTTGGCAGCCACTTCCATGTCTTCCCTCC 1570
Qy      1423 TGAAGAGTCAAGCACCCTC-----CAGTTGACACAGGTATGGGCTTCCACAGATTA 1476
Db      1571 TCACAGTCAAGCACCCTCAGTGTCTTACTCAGACACAGGTGTGGGCTTCCGACAGTA 1630
Qy      1477 CCAAGAAATCTGTCAATGAAAGAAAGAACTGTGAGTTTAACTTGACATTTCCAGAGAG 1536

```

```

Db      1631 CCAAGAGCCTGTCACTGAAGAAAGAAACAGTGAAGTTTAACTTGACATTTCCAGAGTG 1690
Qy      1537 CCCACAGAACATCTTGAACAGCGCGCTTTGACAGATGTCACATATATACAGGCTCT 1596
Db      1691 CCTACGGAACA---TGAACAACCCCATTTGGACCAAGTG-----CACACAT 1735
Qy      1597 AATGAAGAGCATGACAGAGTCTTATCTTACAGGAGTGAAGAGTGGACCAAACTGTCTT 1656
Db      1736 CTTGAAGAGGCTGAGAGATCTTATCTGTCAGATGATTAAGACTCAGACCAATCTGTTTC 1795
Qy      1657 CTATGATGATGCTTCAACAGCTGGCCCTGACTTGCAAGATGCCACAGGAGCATGAC 1716
Db      1796 CTATGATGATGCTTCAACAGCTGGTTCTCACTGCAATGATG-CAAGCAGGAGCATGAC 1854
Qy      1717 CCACTCAAGGACACAGAGGAGCACCAGAGAGTGTTCACAGCAGTCTTACAGATG 1776
Db      1855 CCACTCAAGGAGCAGAC-AGCAGCAGACTCAAGAAATGTTCCATGAGCAGTCTGGCAGG 1913
Qy      1777 TAACTCAAGAACCAAGAGTACTTAATAGTCTGCTGCAAAACACCTGATTTTACAGTCTT 1836
Db      1914 TAACTCAAGAACCAAGAGTCTGCT---TAGTCTGTCGAAACACAGCTGCTTCACTATCTT 1970
Qy      1837 T-CCCAACTPAAGAGTTTAAATAAAGTGAATATTTCTTTTGTGAAGAAAAA 1895
Db      1971 TCCCAAACTPAAGAGTTTAAATAA-----AATATTTCTTTCAGGCAAAAAA 2024
Qy      1896 AAAAAAAAAA 1906
Db      2025 AAAAAAAAAA 2035

RESULT 4
AY357925
LOCUS      2187 bp mRNA linear PRI 26-AUG-2003
DEFINITION Macaca fascicularis beetrophin (VMD2) mRNA, complete cde.
ACCESSION AY357925
VERSION   AY357925.1 GI:34013782
KEYWORDS
SOURCE
ORGANISM  Macaca fascicularis (crab-eating macaque)
           Macaca fascicularis
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
           Cercopithecinae; Macaca.
REFERENCE  1 (bases 1 to 2187)
           Okamoto,H., Umeda,S., Suzuki,M.T., Yoshikawa,Y., Tanaka,Y. and
           Iwata,T.
           Molecular Cloning of VMD2 Gene from Cynomolgus Monkey (Macaca
           fascicularis)
           Unpublished
           2 (bases 1 to 2187)
           Okamoto,H., Umeda,S., Suzuki,M.T., Yoshikawa,Y., Tanaka,Y. and
           Iwata,T.
           Direct Submission
           Submitted (01-AUG-2003) National Tokyo Medical Center, National
           Institute of Sensory Organs, 2-5-1, Higasbisaoka, Meguro-ku, Tokyo
           152-0021, Japan
FEATURES
     source
         1..2187
             /organism="Macaca fascicularis"
             /mol_type="mRNA"
             /db_xref="taxon:9541"
             1..2187
             /gene="VMD2"
             67..1824
             /gene="VMD2"
             /codon_start=1
             /product="beetrophin"
             /protein_id="AA056049.1"
             /db_xref="GI:34013783"
             /translation="MTITITTSQVANNARLGSFSSLLUCRWSITKLYLGEFFILCY
             IIRFIYRLATTEQQLMFEKLTLYCDSYIQILIPISFLGFYVTLVYTRWNOYENLFW
             PDRIMLSVSGFVGKDEGRLLRRTILRYANLGNVLIRSVTAIVYRPFPSAHLVQA
             GFMTPEHKKLEKLSLPNNMFVFWVFAFLSMKAWLGRIRBDIILQSLINENKTLR

```

10GHLVAVDWISIPLYTVVAVVSFFLCLVGRFLNPAKAYRGLHDLVVPE
 TFLQFFVYVWLVKVAEOLINPREGDDDDPENNUTVDNLOVSLAVDEMHDLPNEP
 DMYNBPBPHPTTAAASOPRASPSPKSTFSLINKEMFOPQEDKEDAHGIIGR
 FLGLOSHHHBPGRANSRKLMPKRESLSLHGLPKNKAKOVNVEDKAKMLAV
 DAPKSAPIYORPGTYSAPQPLSPFMEFPPEPEVLSKHSVTVNGIDKDSLKTVSSG
 AKNSSELISDGLAMHEPESVSHVRKXIVFENLIDMBEIPENHLKEPBLESTNHAT
 LKDHVDPMVALENDEAHS"

ORIGIN

Query Match 45.6%; Score 873.8; DB 9; Length 2187;
 Best Local Similarity 77.2%; Pred. No. 2,5e-202;
 Matches 118; Conservative 0; Mismatches 292; Indels 39; Gaps 3;

QY 1 GTGCCAAGCCATGACTATCACTTACACAAACAAAGTGGCCATGCGCGCTGGTTCCT 60
 Db 57 GTGCTGGCCATGACCACTACCTTACAGAGCAAGTGGCTAAAGCCCGCTTAGGCTCTT 116
 QY 61 CTGTCCTCCTCCTGCTGGTGGGAGGACAGATCTTACAGCTGCTGTATGAGAAATTCCT 120
 Db 117 CTCCCGCTGCTGCTGTGGTGGGAGGACAGATCTTACAGCTGCTGTATGAGAAATTCCT 176
 QY 121 TGTCTTCAATTCCTCTACTATTCCATCCGTGAGACTCTACAGATGGTTCCTCGAGTGA 180
 Db 177 CATCTTCCTGCTGCTGCTACTACATCCCTTTATTTATAGGCTGGCCCTCAGCAAGA 236
 QY 181 TCAGCAGCTGTGTGTGAGAAAGCTGCTGTATCTGCAAGCTTACATTCAGCTCATCC 240
 Db 237 GCAACAGCTGATGTGTGAGAAAGCTGCTGTATTCGACAGCTTACATTCAGCTCATCC 296
 QY 241 TATATCCTTCCTGCTGGGTTTCAATGTACATTTGGTGGTGGACCGCTGTGAGGCACTA 300
 Db 297 CATTCCTTCGCTGGTGGCTTCTACGTAGCCGTGGTGGTGGACCGCTGTGAGGCACTA 356
 QY 301 CGAAGACTTGGCGGTCGACCGCTCTCATGATCCAGGTGTCTAGCTGTGAGGCACTA 360
 Db 357 CGAAGACTTGGCGGTCGACCGCTCTCATGATCCAGGTGTCTAGCTGTGAGGCACTA 416
 QY 361 GGATGAGAAAGCCGCTTGTGGGAGGACGCTTACCTCCGCTACGCTATCTGGGCAAGT 420
 Db 417 GGAAGAGCAAGCCCGCTGTGGGAGGACGCTTACCTCCGCTACGCTATCTGGGCAAGT 476
 QY 421 GCTCATCTGGCGACATTCAGCACTCGGTCTTACAGGCTTTCCCACTTTCACCACT 480
 Db 477 GCTCATCTGGCGACATTCAGCACTCGGTCTTACAGGCTTTCCCACTTTCACCACT 536
 QY 481 GGTGTAGAGGTTTATAGACCACTGAGGAACTATGAGAGTTGAGAAAGTGGGCTTACC 540
 Db 537 GGTGTAGAGGTTTATAGACCACTGAGGAACTATGAGAGTTGAGAAAGTGGGCTTACC 596
 QY 541 ACACAACACATTCGTGGTGGCTGGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
 Db 597 ACACAACACATTCGTGGTGGCTGGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 656
 QY 601 TGGAGGTGGAATCCGGGACACCGCTGCTTCAGAGGCTTGAATGAGGTGTGTACTTT 660
 Db 657 TGGAGGTGGAATCCGGGACACCGCTTACCTGCTTCAGAGGCTTGAATGAGGTGTGTACTTT 716
 QY 661 GCGTACTCAGTGTGAGAGCTGTATGCTTCAAGTGTGATAGTATCCCATTTGTGTAC 720
 Db 717 GCGTACTCAGTGTGAGAGCTGTATGCTTCAAGTGTGATAGTATCCCATTTGTGTAC 776
 QY 721 ACAGGTGTGAGAGTGTGAGAGCTTTCCTTGTGATGCTTGTGAGGAGGAGT 780
 Db 777 ACAGGTGTGAGAGTGTGAGAGCTTTCCTTGTGATGCTTGTGAGGAGGAGT 836
 QY 781 TCTGAACCAACAAAGGACTACCAAGGCTTGAATGAGATCTGTGTGCTGTCTTAC 840
 Db 837 TCTGAACCAACAAAGGCTTACCTTGAATGAGATCTGTGTGCTGTCTTAC 896
 QY 841 AATCTGCAATTTCTTATTTACATGAGGCTGTGAGAGTGTGAGAGCAAGCTCATTAACC 900
 Db 897 AATCTGCAATTTCTTATTTATGAGGCTGTGAGAGTGTGAGAGCAAGCTCATTAACC 956

QY 901 CTTCGGGAGAGACATGATGATTTTGAAGCTAATCTGATCATTTAGCAAGAACTGCGAGT 960
 Db 957 GTTTGGAAGAGACATGATGATTTTGAAGCACTGATGATTTGCAAGAAATTTGCAAGT 1016
 QY 961 GTTCCTGTTGTCCTGATGAGATGACACAGAACTTGCCTCCATGAGAAAGTGAATGTA 1020
 Db 1017 GTTCCTGTTGTCCTGATGAGATGATGACACAGAACTTGCCTCCATGAGAAAGTGAATGTA 1076
 QY 1021 CTGGAACAGAGCAGGCTCAGCGCCCTTACAGAGCTTCTTCCAGAGTCTTCCCGGCA 1080
 Db 1077 CTGGAATGAGCTTACGCAACCCCTTACAGAGCTTCTTCCAGAGTCTTCCCGGCA 1136
 QY 1081 TTCTTATGAGGCTTCCACTTCAACATCAGCTTAAAGAAAGAGCTTGAAGCTTGTGTC 1140
 Db 1137 CTCTTATGAGGCTTCCACTTCAACATCAGCTTAAAGAAAGAGCTTGAAGCTTGTGTC 1196
 QY 1141 AAAAGAGAGGCTGACACGATTAAGAAAGAGAGTGTCTTACAGACCATAGGCTGCTT 1200
 Db 1197 CAATCAGAGG-----ACAAAGAGAGCTTCACTGGCATCATTTGGCGCTT 1244
 QY 1201 CTTAGACTGCAACCCAAACCTACCATCTTCCCTTGAAGAGCTTAAAGACCAACTATT 1260
 Db 1245 CTAGAGGCTGCAATCCCATGATCATCTCTCCGAGGCAAACTAAGAACCAACTATT 1304
 QY 1261 GTGTTCTAAGA-----ACCCCTCTGAGAGGCAAGTGAAGATGCCAACAGAAAA 1314
 Db 1305 GTGAGCCAAAGAGGAATCCCTTCTCAGAGGAGGCTTCCCAAAACCAAGGCAAGTCA 1364
 QY 1315 CCAAGAG-----ATGCTGGAATTTAAGGCTGTGAGTCTT 1353
 Db 1365 ACAGAGCTTAGAGGAGCTTGAAGACAAAGAGCTTGAAGCTTGAAGTCTT 1424
 QY 1354 GAAATGTGTTCCAGGTTTAAAGAGAGAGCTCCCATTTGTGGCCACAGGACCCAGAG 1413
 Db 1425 CAGTGTGCCCCAGTATCAGAGGCGCAGCTACTACAGTCCCAAGAGAGGCTTACG 1484
 QY 1414 CCAGCTTAC 1422
 Db 1485 CCCCACTCC 1493

RESULT 5
 AF073501 2171 bp mRNA linear PRI 02-SEP-1998
 LOCUS
 DEFINITION Homo sapiens vitelliform macular dystrophy protein (VMD2) mRNA,
 complete cds.
 ACCESSION AF073501
 VERSION AF073501.1 GI:3511241
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens (human)

REFERENCE
 1 (bases 1 to 2171)
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS
 Stohr,H., Marguardt,A., Rivera,A., Cooper,P.R., Nowak,N.J.,
 Shows,T.B., Gerhard,D.S. and Weber,B.H.

TITLE
 A gene map of the Best's vitelliform macular dystrophy region in
 chromosome 11q12-q13.1

JOURNAL
 MEDLINE
 Genome Res. 8 (1), 48-56 (1998)

REFERENCE
 PUBMED
 9445467
 2 (bases 1 to 2171)
 Marguardt,A., Stohr,H., Paesmore,L., Kraemer,F., Rivera,A. and
 Weber,B.H.F.

TITLE
 JOURNAL
 Submitted (22-JUN-1998) Human Genetics, University, Biozentrum, Am
 Hubland, Wuerzburg 97074, Germany

FEATURES
 source
 1. .2171
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="11"

gene
CDS
/map="11q12-q13.1"
/cfeature.type="retinal pigment epithelium"
1..2171
/gene="VMD2"
89..1846
/gene="VMD2"
/note="Best disease"
/codon.start=1
/product="vitelliform macular dystrophy protein"
/protein.id="AAC33766.1"
/db_xref="GI:3511242"

/translation="MTITVTISQVANARIGSFRLLLCMRGSYKLYGEPLIFELCY
IIRPIRYLALTEBOLEMEKLTJLDCSYIQLIPISFVLGFTVLIVTWMQYENLPM
PDRMLSVSGVEGKDEQGRLLRRLIRYANLGNVLIISVSTAYKRFPAOHLVOA
GFTMPAEHKOLEKSLPHNMFVPMVPMANLMSKMLGRIHDIPLLOSILNMTLR
TOGHLIAYADWISIPIVYTOYVTAAYVSFFLTCLVGRQFAPAKARYBEHLDIYVPR
TFLOPFYVGLKLYABQLINPGEDEDDFFETNMIYDRNLQVSLAVDEMDLPRMER
DMVNNKEPPQPYTAAQAQFRRAFSMSTFNIISLKEBMEKQVPODEEDNAHGIIGR
FLGLOSHDHPBRANSRKLMLPKRESLHLEGLPNHKAQNVRGQEDNKAWLKAV
DAFKSAPLYORPGYSAPOPLSLPTMFPFPLEPSAPLSHSVTGIDTOKSLKTVSSG
AKSFELLESBDGALMEHREVSQVRAKTVENLTDMPRIENHILKEPLKPSPTNIHTT
LKDMDPFWALENDDEANS"

ORIGIN

Query Match 45.6%; Score 872.8; DB 9; Length 2171;
Best Local Similarity 77.1%; Pred. No. 4,4e-202;
Matches 1117, Conservative 0; Mismatches 292; Indels 39; Gaps 3;

QY 2 TGCAGAGCTGACTATATCACTTACAAACAAAGTACCAATGCCCGCTCGGTTCGTTTC 61
DB 80 TGCCTGGCGATGACATCACTTACAAAGTACCAATGCCCGCTCGGTTCGTTTC 139
QY 62 TCGTCCCTCCCTGCTGCTGGGAGGAGCATCTTACAAAGTACCAATGCCCGCTCGGTTC 121
DB 140 TCCCGCTGCTGCTGCTGCTGGGAGGAGCATCTTACAAAGTACCAATGCCCGCTCGGTTC 199
QY 122 GTCTTCAATTCCTCTACTATTCATCCGTGACTCTTACAGATGGTTCCTCGAGTAT 181
DB 200 ATCTTCCGCTCTGCTACTATCATTCGCTTTATTTATAGGCTGGCCCTACGGAAGAA 259
QY 182 CAGAGCTGTTGTTGAGAGCTGCTGTGATCTGCGACAGCTACATTCAGCTCATCCCT 241
DB 260 CAAAGCTGATGTTGAGAAACTGACTGTGATGCGACAGCTACATTCAGCTCATCCCT 319
QY 242 ATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
DB 320 ATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
QY 302 GAGAACTTGCCTGAGCCGAGCCGCTCATGATCCAGTGTCTAGCTTGTGAGAGGCAAG 361
DB 380 GAGAACTTGCCTGAGCCGAGCCGCTCATGAGCTGTGAGGCTTGTGAGAGGCAAG 439
QY 362 GATAGAGAGGCGCTTGTGCTGCGGAGCAGCTCATCGCTACGCCATCTGAGGCAATG 421
DB 440 GACGAGCAAGGCGCGCTGCTGCGGAGCAGCTCATCGCTACGCCAATCTGAGGCAAG 499
QY 422 CTCTATCCGCGCAGCATAGACCTCGGTCTTACAAAGGCTTCCCACTTCCACCACTG 481
DB 500 CTCTATCCGCGCAGCATAGACCTCGGTCTTACAAAGGCTTCCCACTTCCACCACTG 559
QY 482 GTGCTAGAGAGGTTTATGACCCATAGGAGCAATAAGAGCTTGGAGAAATTGAGGCTTACA 541
DB 560 GTGCAAGAGGCTTTATGACTCCGAGCAACAAAGAGCTTGGAGAAACTAGAGCTTACA 619
QY 542 CACAACATTTCTGGGTGCTCGGTGTGTGTTTGCAACTGTGCAATGAGAGGCTTATCTT 601
DB 620 CACAACATTTCTGGGTGCTCGGTGTGTGTTTGCAACTGTGCAATGAGAGGCTTATCTT 679
QY 602 GGAAGTGAATTCGAGGAGACCGTCTGCTCCAGAGCCGAGAGTGAATGAGTGTGATCTTGG 661
DB 680 GGAAGTGAATTCGAGGAGACCGTCTGCTCCAGAGCCGAGAGTGAATGAGTGTGATCTTGG 739
QY 662 CGTACTCAGTGTGACAGCTGTATGCTTACGACTGATGATGATTCATTCATGCTGTGATACA 721

DB 740 CGTACTCAGTGTGACAGCTGTATGCTTACGACTGATGATGATTCATTCATGCTGTGATACA 799
QY 722 CAGGTGTGACAGTGTGAGAGTATACAGCTTTTCCCTGCAATCTGATGGAGGAGCTTT 781
DB 800 CAGGTGTGACTGTGGCGGTGTACAGCTTCTTCCGATTTCTATGTTGGGAGGAGTTT 859
QY 782 CTGAACCCAAACAAGACTACACAGGAGGAGTATGATGATGATGATGATGATGATGATGAT 841
DB 860 CTGAACCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 919
QY 842 ATCTGCAATCTTATTTCTACATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
DB 920 TTCTGCAATCTTCTTCTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979
QY 902 TTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 961
DB 980 TTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1039
QY 962 TCCCTGTTGCTCGTGAATGGAGTGCACCAAGACTTGCCTCCATGGAAGCTGATGATGAT 1021
DB 1040 TCCCTGTTGCTCGTGAATGGAGTGCACCAAGACTTGCCTCCATGGAAGCTGATGATGAT 1099
QY 1022 TGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1081
DB 1100 TGAATTAAGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1159
QY 1082 TCTTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1141
DB 1160 TCTTTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1219
QY 1142 AAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1201
DB 1220 AATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1267
QY 1202 TTAGAGCTGCAACCCCAAAAATACCATCTTCTTGAAGAATTAAAGCAAACTATTTG 1261
DB 1268 CTAGAGCTGCAAGCTCCATGATCAATCTCTCCAGAGGCAAACTCAAGAGCAAACTATG 1327
QY 1262 TGTCTTAAAG-----ACCCCTCTGGAAGGCGAGTGAAGATGCCAACCAAGAAAAC 1315
DB 1328 TGGGCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1387
QY 1316 CAGAAAG-----ATGCTGGAATTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1354
DB 1388 CAGAACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1447
QY 1355 AAATGTGTTCCAGGTTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1414
DB 1448 AAGTCTGCCCACTGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1507
QY 1415 CACCTTAC 1422
DB 1508 CCCACTCC 1515

RESULT 6
AF057169 2210 bp mRNA linear PRI 17-OCT-1998
LOCUS Homo sapiens beetrophin (VMD2) mRNA, alternatively spliced product,
DEFINITION complete cds.
ACCESSION AF057169
VERSION AF057169.1 GI:3335158
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Beryozoa; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2210)
Petrukhin, K., Koishi, M., J., Bakall, B., Li, W., Xie, G., Matkovic, T.,
Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M.,
Bergan, A. A., McGarty-Dugan, V., Figueroa, D., Austin, C. P.,
Metzker, M. L., Caskey, C. T. and Madeline, C.

TITLE Identification of the gene responsible for Best macular dystrophy
 JOURNAL Nat. Genet. 19 (3), 241-247 (1998)
 MEDLINE 98324772
 PUBMED 9662395
 REFERENCE 2 (bases 1 to 2210)
 AUTHORS Petrukhin, K.
 TITLE Direct Submission
 JOURNAL Submitted (03-APR-1998) Human Genetics, Merck Research
 Laboratories, West Point, PA 19486, USA
 FEATURES
 source
 1. .2210
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 1. .2210
 /gene="VMD2"
 105. .1862
 /gene="VMD2"
 /note="Best's macular dystrophy gene; BMD; alternatively
 splice isoform"
 /product="bestrophin"
 /codon_start=1
 /protein_id="AA064343.1"
 /db_xref="GI:335159"
 /translation="MTITVTSQVANAARLGSFSRLLLCWRGSIYKLLYGEPLIFLCYY
 IIRFYRLALTEBOQIMFEKLTLYCDSYIQLPISEFVIGYVTVLVTVTMNOYENLP
 PRMLSLVSGFVGEKDEQGRLLRLTILRYANLGNVLLRSVSTAYKRPFAHAYOA
 GMPATPEHKELEKSLPHNMFVPMWPMVPMANISMKAMIGLRIDPILLOSLENMTL
 TOCGHLAYDWISILPLYOVTVAVSPEFLICVGRPLNPAKAYBEHELDLVVPR
 TLOPFYVYGVKVAEOLINPFGEDDDDEFTWIVDRNLQVSLAVIDEHLDPMEP
 DMVANKPEPQPYTAASQPRFASFMGSTFNISLNKEMEQPOVEDEDHAGIIGR
 FGLQSHDHPPRANSRTKLMPKRESLLHBEGLPNHNAKAQNVGQEDNDAKMLKAV
 DKFKAPLYORPGYASAPDTPLSPTMFEFLEPSASLSLHSGTIGDTDKSLKTVSSG
 AKSFPELLSESDGALMEHEVSVQRKTVENFLTMPEIIPENHLKEPLEGSPNINHTT
 LKHMDPYALENRDEAHS"

ORIGIN

Query Match 45.5%; Score 871.2; DB 9; Length 2210;
 Best Local Similarity 77.1%; Pred. No. 1.1e-201;
 Matches 1116; Conservative 0; Mismatches 293; Indels 39; Gaps 3;

```

QY      2  TCCCAAGCATGACATCATCACTACCAAAAGATGACCAATGCCCGCTGGTTCGTTCC 61
DB      96  TCCCTGGCATGACATCATCACTACCAAGGCTGTAATGCCCGCTGGTTCGTTCC 155
QY      62  TCGTCCCTCTCTCTGTGTGTGGGAGGAGCATCTACAGAGCTGTGTATGAGAAATTCCT 121
DB      156  TCCCGCTGTGTGTGTGTGTGGGAGGAGCATCTACAGAGCTGTGTATGAGAAATTCCT 215
QY      122  GTCTTCATATTCCTCTATCTATTCATCCGTGAGCTCTACAGAAATGTTCTCTGAGTGAT 181
DB      216  ATCTTCCTGTCTGTCTACTACATATCCGCTTAATTTATAGGCTGGCCCTCAAGAA 275
QY      182  CAGAGCTGTGTGTGAGAACTGTCTGTATGCGACAGACTACATTCAGCTCATCCCT 241
DB      276  CAACAGCTGATGTGTGAAACTGTCTGTATGGACAGCTACATCCAGCTCATCCCT 335
QY      242  ATATCTCTCTGTGTGGGTCTCTAATGTATCAATGGTGTAGCCGCTGTGAGCAAGTAC 301
DB      336  ATTTCTCTCTGTGTGGGTCTCTAATGTATCAATGGTGTAGCCGCTGTGAGCAAGTAC 395
QY      302  GAGAACTTGGCCGTGGCCGACCGCTCATGTACAGGTGTCTGCTTGTGAGAGGCAAG 361
DB      396  GAGAACTTGGCCGTGGCCGACCGCTCATGTAGCTGTGTGTGAGAGGCAAG 455
QY      362  GATGAGAAAGCGCTTTGTCTGTGGCGGACGCTCATCGCTACGCAATCCCTGGGCAAGT 421
DB      456  GACGAGAAAGCGGCTGTGTGGCGGACGCTCATCGCTACGCAATCCCTGGGCAAGT 515
QY      422  CTGATCTGTGGAGCATCAGACCTGTGCTAACAAGCTTTCCACTTTACCAACCTG 481
          |||||

```

```

DB      516  CTGATCTGTGGAGCGTCAGACCGCAGTCTACAGGCGTTCCCGAGGCCAGACCTG 575
QY      482  GTGTAGCAGGTTTTATGACCCATGAGGGAACATAAGAGTTGCGAAGTTGGCTTACCA 541
DB      576  GTGCAAGAGGCTTTATGACTCCGACAGAACAGAGATTTGAGAACTGACCTTACCA 635
QY      542  CACAACATCTTGTGGTGGCCCTGGGTGTGGTTGGCCCACTTGTCAATGAGAGCTTACTT 601
DB      636  CACAACATGTCTGGGTGGCCCTGGGTGTGGTTGGCCCACTGTCTAATGAGAGGCTT 695
QY      602  GGAGGTGAATCCGGGACACCGCTCTGCTCAAGAGCTGTATGATAGGTGTACTTTG 661
DB      696  GGAAGTGAATCCGGGACACCGCTCTGCTCAAGAGCTGTATGATAGGTGTACTTTG 755
QY      662  CGTACTAGTGTGACAGCTGTATGCTTACAGCTGTATGATAGGTGTACTTTG 721
DB      756  CGTACTAGTGTGACAGCTGTATGCTTACAGCTGTATGATAGGTGTACTTTG 815
QY      722  CAGGTGTGACAGTGTGAGTATACAGCTTTTCTTGTGATGCTGTATGCGGAGGAGTT 781
DB      816  CAGGTGTGACAGTGTGAGTATACAGCTTTTCTTGTGATGCTGTATGCGGAGGAGTT 875
QY      782  CTGAACCCAAACAAGACTTACCCAGGCTATGAGATGATCTGTTGTGCTTTCACA 841
DB      876  CTGAACCCAAACAAGGCTTACCCAGGCTATGAGATGATCTGTTGTGCTTTCACA 935
QY      842  ATCTCGAATTTCTTATCTTACATGGGTGTGAAAGTGTGAGAAACAGCTCATCAACC 901
DB      936  TTCTCGAATTTCTTATCTTATGCTGTGCTGAAAGTGTGAGAAACAGCTCATCAACC 995
QY      902  TTGGGAGAGAGATGATATTTTGAAGTACTACCTGATCATTTGACAGAAACCTGAGGT 961
DB      996  TTGGGAGAGAGATGATATTTTGAAGTACTACCTGATCATTTGACAGAAATTTGACAGGT 1055
QY      962  TCCCTGTGTCCGTGATGAGATGACAGAACTTGCCTCCATGGAAGTGAAGTATGAT 1021
DB      1056  TCCCTGTGTGTGATGATGATGACAGAGACTGCTCGGATGGAACCGGACATGAT 1115
QY      1022  TGGAAAGAGAGAGGCTTACGCGGCTTACACAGTGTCTTCTGCCAGGTCTGCGCGCAT 1081
DB      1116  TGGAAAGAGAGAGGCTTACGCGGCTTACACAGTGTCTTCTGCCAGGTCTGCGCGCAT 1175
QY      1082  TCTTCATGAGGCTTCACTTCAATCACTGCTTAAGAAAGAACTTAAGAGCTTGTGCA 1141
DB      1176  TCTTCATGAGGCTTCACTTCAATCACTGCTTGAAGAAAGAGAGATGAGATTCAGGCC 1235
QY      1142  AAAGAGAGGCTGACACGGAATPAAGAAAGAGTGTATAGCAGCAGCATAGTGTCTT 1201
DB      1236  AATCAGAGAGAGAGAGAGATGCTCAGGCTG-----GCATCATGGCGGCTT 1283
QY      1202  TTAGAGCTGCAACCAAAATACCATCTTCTTGAAGCTTAAGAAAGAACTTATG 1261
DB      1284  CTAGGCTGTGAGTCCCATGATGATCATCTCTCCAGGAGCAATCTAAGAGCAAACTAC 1343
QY      1262  TGTTTTAAAG-----ACCCCTCTTGAAGAGCGGTATGAGATGCAACCAAGAAAC 1315
DB      1344  TGGCCCAAGAGAGAAATCTTCTCAGAGGCGCTTGCACAAACCAAGAGAGCAAA 1403
QY      1316  CAGAAAG-----ATGTGTGAATAATTTAAGGCTGTGAGCTTCTT 1354
DB      1404  CAGAAAGCTTGAAGGAGCGCAGAGAAACAAAGAGCTTGAAGGCTGTGAGAGCGCTT 1463
QY      1355  AATGTGTTCAGAGTTTAAAGAGAGAGGCTCCATTTGTGAGCCCAAGGACCCAGCAG 1414
DB      1464  AAGTGTGAGCTGATGATGAGAGGAGGCTACTTACAGTGTGAGAGAGAGAGAGGCT 1523
QY      1415  CACCTTAC 1422
DB      1524  CCACTTCC 1531

```

RESULT 7
 BD136720

LOCUS BD136720 2229 bp DNA linear PAT 18-SEP-2002
DEFINITION Best's macular dystrophy gene.
ACCESSION BD136720
VERSION BD136720.1 GI:23231665
KEYWORDS JP 2002504559-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 2229)
Petrushin,K., Caskey,T.C., Metzker,M. and Wadelius,C.
Best's macular dystrophy gene
Patent: JP 2002504559-A 2 12-FEB-2002;
MERCK & CO INC, CLAS MADELIOUS
COMMENT OS Homo sapiens (human)
PN JP 2002504559-A/2
PD 12-FEB-2002
PF 12-FEB-1999 JP 2000533447
PR 25-FEB-1998 US 60/075941,18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUSHIN,THOMAS C CASKEY,MICHAEL METZKER,CLAS PI
MADELIOUS
PC C07K16/18,C07K14/47,C12N5/10,C12N15/09,C12P19/34,C12Q1/68// PC
C12P21/08,
PC C12N5/00,C12N15/00
CC Best's macular dystrophy gene
FH Key Location/Qualifiers
FT source 1..2229
/organism="Homo sapiens (human)".
/location="Qualifiers"
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 45.5%; Score 871.2; DB 6; Length 2229;
Best Local Similarity 77.1%; Pred. No. 1,1e-201;
Matches 1116; Conservative 0; Mismatches 293; Indels 39; Gaps 3;
Query 2 TGCAGGCGATGACTATCATCACTACCAAAAGAGGCAATGCCGCTCGGTTGCTTC 61
Db TGCAGGCGATGACTATCATCACTACCAAAAGAGGCAATGCCGCTCGGTTGCTTC 155
Query 62 TCGTCCCTCCCTGCTGCTGGCGAGGAGCATCTTCAAGCTGCTGTATGAGAACTTCTT 121
Db TCGTCCCTCCCTGCTGCTGGCGAGGAGCATCTTCAAGCTGCTGTATGAGAACTTCTT 215
Query 156 TCCGCGCTGCTGCTGCTGGCGAGGAGCATCTTCAAGCTGCTGTATGAGAACTTCTT 215
Db TCCGCGCTGCTGCTGCTGGCGAGGAGCATCTTCAAGCTGCTGTATGAGAACTTCTT 275
Query 122 GTCTTCATATTCCTACTATTCATCCGAGACTCTACAGAAATGCTCTCGAGTAT 181
Db ATCTTTCGCTGCTGCTACTATCATCCGCTTATTTATAGGCTGAGCCCTACAGGAA 275
Query 216 ATCTTTCGCTGCTGCTACTATCATCCGCTTATTTATAGGCTGAGCCCTACAGGAA 275
Db ATCTTTCGCTGCTGCTACTATCATCCGCTTATTTATAGGCTGAGCCCTACAGGAA 275
Query 182 CAGAGCTGTTGTTGAGAGCTGCTGTACTGCGACAGCTACATTCACTCATCTCCT 241
Db CAGAGCTGTTGTTGAGAGCTGCTGTACTGCGACAGCTACATTCACTCATCTCCT 335
Query 276 CAAAGCTGATGTTGAGAACTGACTGTATTGCGACAGCTACATCACTCATCTCCTC 335
Db CAAAGCTGATGTTGAGAACTGACTGTATTGCGACAGCTACATCACTCATCTCCTC 395
Query 242 ATATCTTCTGTTGCTGCTTCTATGTTACATTGGTGGAGCCGCTGCTGAGCCAGTAC 301
Db ATATCTTCTGTTGCTGCTTCTATGTTACATTGGTGGAGCCGCTGCTGAGCCAGTAC 395
Query 336 ATTTCTTCTGCTGCTGCTTCTACGTGACGCTGCTGACCCGCTGCTGAGAACCAAGTAC 395
Db ATTTCTTCTGCTGCTGCTTCTACGTGACGCTGCTGACCCGCTGCTGAGAACCAAGTAC 421
Query 302 GAGAACTTGCCTGAGCCGAGCCGCTCATGATCCAGTGTCTAGCTTCTGTGAGGCGAAG 361
Db GAGAACTTGCCTGAGCCGAGCCGCTCATGATCCAGTGTCTAGCTTCTGTGAGGCGAAG 455
Query 396 GAGAACTTGCCTGAGCCGAGCCGCTCATGATCCAGTGTCTAGCTTCTGTGAGGCGAAG 455
Db GAGAACTTGCCTGAGCCGAGCCGCTCATGATCCAGTGTCTAGCTTCTGTGAGGCGAAG 515
Query 456 GAGAGCAAGGCGCTGCTGCTGCGGCGACGCTCATCCGCTACGCCAACCTGAGCAAGCTG 515
Db GAGAGCAAGGCGCTGCTGCTGCGGCGACGCTCATCCGCTACGCCAACCTGAGCAAGCTG 541
Query 422 CTATCTCTGCGCAGCATAGACCTCGGTCTTCAAGGCGCTTCCCATCTTCCACCACTG 481
Db CTATCTCTGCGCAGCATAGACCTCGGTCTTCAAGGCGCTTCCCATCTTCCACCACTG 575
Query 516 CTATCTCTGCGCAGCATAGACCTCGGTCTTCAAGGCGCTTCCCATCTTCCACCACTG 575
Db CTATCTCTGCGCAGCATAGACCTCGGTCTTCAAGGCGCTTCCCATCTTCCACCACTG 541
Query 482 GTGCTAGAGGCTTTATGACCATAGGAGGAGATAAGAGCTTGGAGCAAGTTGGGCTTACCA 541
Db GTGCTAGAGGCTTTATGACCATAGGAGGAGATAAGAGCTTGGAGCAAGTTGGGCTTACCA 541

Db 576 GTGCAAGAGGCTTTATGACTCCGCGAGAACACAGCAGTTGGAGAACTGAGCTTACCA 635
Query 542 CACAACATCTCTGAGGCGCTGAGGCTGTTGCTTCCAACTTGTAAATGAGGCCATCTT 601
Db CACAACATCTCTGAGGCGCTGAGGCTGTTGCTTCCAACTTGTAAATGAGGCCATCTT 695
Query 602 GAGAGTGAATCCCGGAGCACCGCTGCTCCAGAGCCGTATGAAATGAGGTGTACTTTG 661
Db GAGAGTGAATCCCGGAGCACCGCTGCTCCAGAGCCGTATGAAATGAGGTGTACTTTG 755
Query 662 GTACTCAGTGTGACAGCTGTATGCTTACAGCTGATGATGATCCATGGTGTACCA 721
Db GTACTCAGTGTGACAGCTGTATGCTTACAGCTGATGATGATCCATGGTGTATACA 815
Query 722 CAGGTGTGACAGTGTATGAGCTTTTCTGTGATGCTTGTATGAGGAGGAGTT 781
Db CAGGTGTGACAGTGTATGAGCTTTTCTGTGATGCTTGTATGAGGAGGAGTT 875
Query 816 CAGGTGTGACAGTGTATGAGCTTTTCTGTGATGCTTGTATGAGGAGGAGTT 875
Db CAGGTGTGACAGTGTATGAGCTTTTCTGTGATGCTTGTATGAGGAGGAGTT 901
Query 842 ATCTGCAATCTTATTTCTACATGAGGCTGCTGAGAGGTGAGAACTGATCAACCCC 901
Db ATCTGCAATCTTATTTCTACATGAGGCTGCTGAGAGGTGAGAACTGATCAACCCC 995
Query 936 TTCTGCGAGTTCTTCTTATGTTGCTGCTGAGAGGTGAGAGAGCTCATCAACCCC 995
Db TTCTGCGAGTTCTTCTTATGTTGCTGCTGAGAGGTGAGAGAGCTCATCAACCCC 961
Query 902 TTGCGGAGAGCATGATGATTTTGAAGCTAACTGATCATTTGACGAAACCTGCAAGTG 961
Db TTGCGGAGAGCATGATGATTTTGAAGCTAACTGATCATTTGACGAAACCTGCAAGTG 1055
Query 996 TTTGAGAGAGTATGATGATTTTGAAGCAACTGATTTGACAGAGAAATTTGCAAGTG 1055
Db TTTGAGAGAGTATGATGATTTTGAAGCAACTGATTTGACAGAGAAATTTGCAAGTG 1021
Query 962 TCCGCTGTTGCTGAGTATGAGTATGACCAAGAACTTGCCTCCATGAGAACTGATGATAC 1021
Db TCCGCTGTTGCTGAGTATGAGTATGACCAAGAACTTGCCTCCATGAGAACTGATGATAC 1115
Query 1056 TCCGCTGTTGCTGAGTATGAGTATGACCAAGAACTTGCCTCCATGAGAACTGATGATAC 1115
Db TCCGCTGTTGCTGAGTATGAGTATGACCAAGAACTTGCCTCCATGAGAACTGATGATAC 1175
Query 1022 TGAAGCAGAGGAGGCTCAGCGCTTACACAGCTGCTTGCAGAGTCTGCGGAGT 1081
Db TGAAGCAGAGGAGGCTCAGCGCTTACACAGCTGCTTGCAGAGTCTGCGGAGT 1175
Query 1116 TGAAGTAAAGCCGAGGAGCCAGCCCTTACACAGCTGCTTGCAGAGTCTGCGGAGT 1175
Db TGAAGTAAAGCCGAGGAGCCAGCCCTTACACAGCTGCTTGCAGAGTCTGCGGAGT 1141
Query 1082 TCTTTCATGAGCTTCACTTCAATCACTGAGCTTAAAGAAAGACTTGAAGCTTGTGCA 1141
Db TCTTTCATGAGCTTCACTTCAATCACTGAGCTTAAAGAAAGACTTGAAGCTTGTGCA 1235
Query 1176 TCTTTCATGAGCTTCACTTCAATCACTGAGCTTAAAGAAAGACTTGAAGCTTGTGCA 1235
Db TCTTTCATGAGCTTCACTTCAATCACTGAGCTTAAAGAAAGACTTGAAGCTTGTGCA 1201
Query 1142 AAAGAGAGGCTGACAGGATGAGAAAGAGAGTGTATGAGCAAGCTTGAAGCTTGTGCA 1201
Db AAAGAGAGGCTGACAGGATGAGAAAGAGAGTGTATGAGCAAGCTTGAAGCTTGTGCA 1283
Query 1236 AATGAGAGAGCGAGAGAGTGTATGAGCAAGCTTGAAGCTTGTGCA 1283
Db AATGAGAGAGCGAGAGAGTGTATGAGCAAGCTTGAAGCTTGTGCA 1261
Query 1202 TTAGAGCTGCAACCAAAATCACTTCTTCTTGAAGAACTTAAAGCAAACTATATG 1261
Db TTAGAGCTGCAACCAAAATCACTTCTTCTTGAAGAACTTAAAGCAAACTATATG 1343
Query 1284 CTAGGCTGAGCTTCAATGATCACTTCTTCAAGGAGAACTCAAGGCAAACTATATG 1343
Db CTAGGCTGAGCTTCAATGATCACTTCTTCAAGGAGAACTCAAGGCAAACTATATG 1315
Query 1262 TGTTCATAGA-----ACCCCTCTTGAAGGCGAGTGAAGATGCCAACCAAGAAAC 1315
Db TGTTCATAGA-----ACCCCTCTTGAAGGCGAGTGAAGATGCCAACCAAGAAAC 1403
Query 1344 TGGCCCAAGAGGAAATCCTTCTCAGAGGAGGCTTGCACCAAGGAGGAGGAGGAGG 1403
Db TGGCCCAAGAGGAAATCCTTCTCAGAGGAGGCTTGCACCAAGGAGGAGGAGGAGG 1354
Query 1316 CAGAAAG-----ATGCTGAAATTTAAGGAGTGTGAGCTTGTG 1354
Db CAGAAAG-----ATGCTGAAATTTAAGGAGTGTGAGCTTGTGAGCTTGTG 1463
Query 1404 CAGAACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1463
Db CAGAACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1414
Query 1355 AATATGTTTCCAGAGTTTAAAGAGAGGCTTCCATGTTGAGCCCAAGGAGCCAGGAGC 1414
Db AATATGTTTCCAGAGTTTAAAGAGAGGCTTCCATGTTGAGCCCAAGGAGCCAGGAGC 1523
Query 1464 AAGTCTGAGCCAGTATCAGAGGCGAGGCTTACAGTGTGCTTCCAGAGGAGGAGGAGG 1523
Db AAGTCTGAGCCAGTATCAGAGGCGAGGCTTACAGTGTGCTTCCAGAGGAGGAGGAGG 1531
Query 1415 CACCTTAC 1422
Db CACCTTAC 1531

RESULT 8
AX745964 1758 bp DNA linear PAT 14-MAY-2003
LOCUS AX745964
DEFINITION Sequence 1 from Patent WO03030922.
ACCESSION AX745964
VERSION AX745964.1 GI:30724619

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 Steuernagel, A., Bruenner, G., Fritsch, R., Eulenberger, K. and Closek, T.

TITLE Bestrophin and bestrophin homologous proteins involved in the regulation of energy homeostasis

JOURNAL Patent: WO 03030922-A, 17-APR-2003; Develcoen Aktiengesellschaft fuer Entwicklungsbiologische Forschung (DE)

FEATURES

source Location/Qualifiers

1..1758

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 45.3%; Score 867; DB 6; Length 1758;

Best Local Similarity 77.1%; Pred. No. 1.1e-200;

Matches 1110; Conservative 0; Mismatches 290; Indels 39; Gaps 3;

QY 11 ATGACTATCACTACACAAACAAAGTACCAATGCCCGCTGCTTCTTCCTCCCTC 70

Db 1 ATGACCATCACTTACACAGCCAAAGTGTATATGCGCTTACGGCTCTTCCGCGCTG 60

QY 71 CTCTGTGTGCTGGGAGGACATCTACAAAGTGTATATGAGAAATTCCTTCTTCA 130

Db 61 CTCTGTGTGCTGGGAGGACATCTACAAAGTGTATATGAGAAATTCCTTCTTCA 120

QY 131 TTCTCTACTATTCCTTCCGTGACCTCTACAGATGTCTTCAAGTATCAGAGCTG 190

Db 121 CTCTGTACTATTCCTTCCGTGACCTCTACAGATGTCTTCAAGTATCAGAGCTG 180

QY 191 TTGTGTAGAAAGTGTGCTGTACTGACAGCTACATTCAGCTTACCTTATCCCTC 250

Db 181 ATGTGTAGAAAGTGTGCTGTACTGACAGCTTATGAGCTTACCTTATCCCTC 240

QY 251 GTTCTGGGTTTCTATGTTTCAATTGTGTGTGAGCCGCTGTGTAGAGGAGAGAGAA 310

Db 241 GTTCTGGGTTTCTATGTTTCAATTGTGTGTGAGCCGCTGTGTAGAGGAGAGAA 300

QY 311 CGGTGGGCGGACCGGCTCATGATCCAGGTGTCTAGCTTGTGTAGAGGAGAGAGAA 370

Db 301 CGGTGGGCGGACCGGCTCATGATCCAGGTGTCTAGCTTGTGTAGAGGAGAGAA 360

QY 371 GGCCGCTTGTGCGGCGGACGCTCATCGGTAACGCTGAGGCAAGTGTCTATCTG 430

Db 361 GGCCGCTTGTGCGGCGGACGCTCATCGGTAACGCTGAGGCAAGTGTCTATCTG 420

QY 431 CGGACATCAGACCTCGGTCTACAAAGGCTTCCCACTCTTACCACTGTGTGAGCA 490

Db 421 CGGACGTCAGACCTCGGTCTACAAAGGCTTCCCACTCTTACCACTGTGTGAGCA 480

QY 491 GGTTTTATGACCACTGGGAAACATAGCAAGTTGAGAAAGTTGGGCTTACCAACA 550

Db 481 GGTTTTATGACCTGGGAAACATAGCAAGTTGAGAAAGTTGGGCTTACCAACA 540

QY 551 TTCTGGGAGCCCTGGGTGTGTGTGTTCCAACTTGTCAATGAAGGCTTATCTTGA 610

Db 541 TTCTGGGAGCCCTGGGTGTGTGTGTTCCAACTTGTCAATGAAGGCTTATCTTGA 600

QY 611 ATCCGGGACACCGTCTGTCTCAAGAGCTGATGAATGAGGTGTGTCTTGGTACTG 670

Db 601 ATCCGGGACACCGTCTGTCTCAAGAGCTGATGAATGAGGTGTGTCTTGGTACTG 660

QY 671 TGTGAGAGCTGTATGCTTACAGTGTATGATATCCATGTGTGTACACAGGTGTG 730

Db 661 TGTGAGAGCTGTATGCTTACAGTGTATGATATCCATGTGTGTACACAGGTGTG 720

QY 731 ACAGTGGACATATACAGCTTTTCTTGTGATGCTTATGAGGAGGAGATTTCTGA 790

Db 721 ACTGTGGCGGTGTACAGCTTCTTCTGACTGTGTGTGTGGGCGGAGTTCTAACCCA 780

QY 791 AACAGGACTACCCAGGCGCATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 850

Db 781 GCCAAGGCGCTTACCCAGGCGCATGATGTGTGTGTGTGTGTGTGTGTGTGTGT 840

QY 851 TTCTTATTTCTACATGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 910

Db 841 TTCTTCTTCTATGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900

QY 911 GACGATGTATTTTGT 970

Db 901 GATGATGTATTTTGT 960

QY 971 TCCGTGATGT 1030

Db 961 GCTGT 1020

QY 1031 GCAGGCGCTTACGCGGCTTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1090

Db 1021 CCCAGGCGCAGGCGGCTTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080

QY 1091 GGTCTCACCTTCAACATCAGGCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1150

Db 1081 GGTCTCACCTTCAACATCAGGCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1140

QY 1151 GCTGACAGGATTAAGAAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1210

Db 1141 G-----ACAGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1188

QY 1211 CAACCCAAACATCAGGCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1270

Db 1189 CAGTCCAGATATCAGGCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1248

QY 1271 A-----ACCCCTCTCTCAAGGCGCAGTGTGTGTGTGTGTGTGTGTGTGTGT 1332

Db 1249 AGGAAATCCCTTCTCAGAGGCGGCTGCAAAACCAAGGAGCCCAAGAGAGGTT 1308

QY 1323 -----ATGTGTGAATTTAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1363

Db 1309 AGGAGCGCAGAAAGAAAGAGGCTTAAAGGCTGTGTGTGTGTGTGTGTGTGT 1368

QY 1364 CCAAGGTTTAAAGAGAGGCTCCCATTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1422

Db 1369 CCAGTGTATCAGAGGCGAGGCTTATCAGTGTGTGTGTGTGTGTGTGTGTGTGT 1427

RESULT 9

AY515704 1758 bp mRNA linear PRI 28-JAN-2004

LOCUS AY515704

DEFINITION Homo sapiens bestrophin 1 mRNA, complete cds.

ACCESSION AY515704

VERSION AY515704.1 GI:41216872

KEYWORDS

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE

1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

2 Sun, H., Tsunenari, T., Yau, K.W., and Nathans, J.

3 The vitelliform macular dystrophy protein defines a new family of chloride channels

4 Proc. Natl. Acad. Sci. U.S.A. 99 (6), 4008-4013 (2002)

5 11904445

6 2 (bases 1 to 1758)

7 Tsunenari, T., Sun, H., Williams, J., Cahill, H., Smallwood, P., Yau, K.W., and Nathans, J.

8 Structure-function analysis of the bestrophin family of anion channels

9 JOURNAL U. Biol. Chem. 278 (42), 41114-41125 (2003)

10 MEDLINE 22917481

PUBMED 12907679
3 (bases 1 to 1758)
REFERENCE Sun.H., Taunenari,T., Yau,K.-W. and Nathans,J.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (02-JAN-2004) Molecular Biology, Neuroscience, Johns
Hopkins University, 725 N. Wolfe Street, Baltimore, MD 21205, USA
FEATURES
source
1..1758
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..1758
/note="chloride channel"
/codon_start=1
/product="bestrophin 1"
/protein_id="AAR9654.1"
/db_xref="GI:41216873"
/translation="MTITVTSQVANARLGSFRLLLCWRGSLTKLYGEFLIFLCYV
IIRFVRLALTEBOQLMEKLTLYCDSYIQLIPISFVLGPFVTLVTTVMNQLP
PDRMLSLVSGVEGDEOGRLLRLIRLVANGLILRSVSTAVYKRPASOHLVOA
GPMTPAEHKKOLESLIPNMFVMPVMPFANISMKAMGRIRADPILLOSILNMTLR
TQCHLYAVDWISLIPVYTVYVAVVSFPLTCLVGRPLPAKAYRPEHLDVYVPR
TFLOPFYVGMVKVAEQLNPFGEEDDSEFTNMLVDRNLQVSLAVDEHODLPRMEP
DMYNKEPQPPYTAASQFRRASFMGTFNLSLKEEMEPQPOBEDDAAG1IGR
FLGQSDHHPRANSTRKLLMPKRESLHEGLPKYHKAACNVRGQEDNKAHLKAV
DAFSAPLVORPGYVAPQTPLSPTMPFPLEPSPAPSLHSVTGIDTOKSLKTVSSG
AKSPELISESDGALMEHREVSQVRKRTVERNLDMPEIPEHNLKEPLEQPTNIHTT
LKDMDPFWALENDENAS"

ORIGIN
Query Match 45.2%; Score 865.4; DB 9; Length 1758;
Best Local Similarity 77.1%; Pred. 2.8e-200;
Matches 1109; Conservative 0; Mismatches 291; Indels 39; Gaps 3;

Qy 11 ATGACTATACCTACACAAACAAAGTACCAATGCCCTCGCTTCCTTCCTCCCTC 70
Db 1 ATGACCATCACTTACACAAAGCAAGTGCCTAATGCCCTTACGCTCCCGCCTG 60
Qy 71 CTCCTGCTGCTGCGGAGAGCATCTACAAAGCTGCTATGAGAAATTCCTTCTTCA 130
Db 61 CTGCTGCTGCTGCGGAGAGCATCTACAAAGCTGCTATGAGAAATTCCTTCTTCTG 120
Qy 131 TTCCTCTACTATTCCTGCTGAGACTCTACAGAAATGTTCTCTCGAGTACAGAGCTG 190
Db 121 CTCCTGCTACTATTCCTGCTGAGACTCTCTTATTAAGGCTGCGCCCTCAGGAAAGAACAGCTG 180
Qy 191 TTGTTTGAAGAGCTGCTCTGTACTGCGACAGCTACATTCAGCTCATCTTATCTTC 250
Db 181 ATGTTTGAAGAACTGACTCTGTATTGCGACAGCTACATTCAGCTCATCTCCATTTCTTC 240
Qy 251 GTTCTGGGTTTCTATGTTTACATTTGGTGGAGCCGCTGCTGAGGCAATGAGAACTTG 310
Db 241 GTGCTGGGCTTCTATGATACGCTGCTGAGCCGCTGCTGAGAACAAATGAGAACTTG 300
Qy 311 CCGTGGCCCGACCGGCTCATGATCCAGGATGCTAGCTCTGAGAGGAGGAGGAGTAA 370
Db 301 CCGTGGCCCGACCGGCTCATGAGCTGATGCTGCTGAGAGGAGGAGGAGGAGGAG 360
Qy 371 GACCGTTTCTGCTGCGGAGAGCTCATCCGCTACCGCATCTGAGGCAAGTGTCTACTG 430
Db 361 GACCGGCTGCTGCGGAGAGCTCATCCGCTACCGCATCTGAGGCAAGTGTCTACTG 420
Qy 431 CGCAGCATCAGCACTCGGCTCTACCAAGGCTTTTCCCACTCTTACCACTGTGTGTA 490
Db 421 CGCAGCGTACACCGCAGCTCTACCAAGGCTTTCCCAAGGCTGCTGTAACCA 480
Qy 491 GATTATATGACCATGAGGAGAACTAAGAGGTTGAGAGGTTGAGGCTTACCAACACA 550
Db 481 GACTTTATGACTCCGCGAGAAACAGAGGTTGAGGAACTGAGGCTTACCAACACA 540
Qy 551 TTCTGGAGTCCCTGGGCTGTGTTTCCCACTTGTCAATGAAGGCTTATCTTGGAGTGA 610
Db 541 TTCTGGAGTCCCTGGGCTGTGTTTCCCACTTGTCAATGAAGGCTTGTGAGTGA 600

Qy 611 ATCCGGACACCGCTGCTGCTCCAGAGCCTGATGATGAGTGTGTACTTTGGCTACTG 670
Db 601 ATCCGGACCGCTATCTGCTGCTCCAGAGCCTGCTGAGAGAGATGAACACTTGGTACTG 660
Qy 671 TGTGACAGCTGTATGCTTACAGCTGATGATGATATCCCATTTGTGTACACAGGTGTG 730
Db 661 TGTGACAGCTGTATGCTTACAGCTGATGATGATATCCCATTTGTGTACACAGGTGTG 720
Qy 731 ACAGTGGATATACAGCTTTTCTTCCATGCTGATCCGGAGGCACTTTCTGAACCA 790
Db 721 ACTGTGGGAGTACAGCTTTCTTCCATGCTGATGCTGAGGCACTTTCTGAACCA 780
Qy 791 AACAGAGCTACCCAGGCGCATGATGATGATGCTGTGTGCTGCTTACATCTGCA 850
Db 781 GCCAAGGCTTACCTGCGCATGAGCTGACCTGCTGTGCGCTTCACTGCTGAG 840
Qy 851 TTCTTATTTATCATGAGCTGCTGAGAGTGGACAGAACGCTCATCAACCCCTTGGGAG 910
Db 841 TTCTTCTTCTATGTTGGCTGCTGAAGGTGACAGAGAGCTCATCAACCCCTTGGAGAG 900
Qy 911 GACGATGATGATTTTGAAGTACTAATGATCATTTGACAGAAACCTGCAGGTGCTCTG 970
Db 901 GATGATGATGATTTTGAAGTACTAATGATCATTTGACAGAAATTTGCAGGTGCTCTG 960
Qy 971 TCCGATGATGATGATGACACAGAACTGCTCCCATGAGACGTGATGATGATGAGAG 1030
Db 961 GCTGTGATGATGATGATGACACAGAACTGCTCCCATGAGAGAGCTGATGATGATGAGAG 1020
Qy 1031 GCAGAGCTTACAGCGCGCTTACACAGCTGCTTCTGCGAGGTCTGCGCGCATTTCTTCA 1090
Db 1021 CCGAGGACACAGCGCGCTTACACAGCTGCTTCTGCGAGGTCTGCGCGCATTTCTTCA 1080
Qy 1091 GCTTCCACCTTCAACATTAAGCTTAAGAAAGAAAGTAAAGCTTGTCTTCAAAAGAG 1150
Db 1081 GCTTCCACCTTCAACATTAAGCTTAAGAAAGAAAGTAAAGCTTGTCTTCAAAAGAG 1140
Qy 1151 GCTGACAGGATTAAGAAAGAGTGGCTATGACAGACCATGAGCTGCTTCTTGAAGT 1210
Db 1141 G-----ACGAGAGAGATGCTACCGCTGATCATTTGCGCGCTTCTGAGGCTG 1188
Qy 1211 CAACCCAAATACTACATCTTCTTGAAGACTTAAAGCAATTAATGTGTTCTAAG 1270
Db 1189 CAGTCCCATGATCAACATCTTCCAGGCAAACTCAAGAGCAAACTACTGTGGCCAAAG 1248
Qy 1271 A-----ACCCCTCTCTGAAAGGCCAGTGTAAAGATGCCAACAGAAAACCAAGAA-- 1322
Db 1249 AGGAAATCCCTTCTTCCAGAGAGGCTGCTCCAAAACCAAGGAGCCCAACAGAACGTT 1308
Qy 1323 -----ATGCTGGAATTTAAAGGCTGGACTTCTTGAATGTGTT 1363
Db 1309 AGGGGCGAGAAACAAACAGAGCTGGAAGCTTAAAGCTGTGAGAGCTTCAAGTCTCC 1368
Qy 1364 CCAAGTTTAAAGAGAGAGGCTCCATTTGTGCGCCACAGGACCCAGAGCAACCTTAC 1422
Db 1369 CCACTGATACAGAGGCCAGCTTACTACAGTCCCAACAGAGCCCTCAGCCCACTCC 1427

RESULT 10
CQ729633 2170 bp DNA linear PAT 03-FEB-2004
LOCUS CQ729633
DEFINITION Sequence 15567 from Patent WO02068579.
ACCESSION CQ729633
VERSION CQ729633.1 GI:42301252
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humaneons or transcripts, for detecting expression and other uses

JOURNAL
Patent: WO 02068579-A 15567 06-SEP-2002;
PR Corporation (NY) (US)

FEATURES
Source
1. .2170
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 44.8%; Score 859.2; DB 6; Length 2170;
Best Local Similarity 77.0%; Pred. No. 9.3e-199;
Matches 1115; Conservative 0; Mismatches 293; Indels 40; Gaps 4;

2 TGGCAAGCCATGATCACTACACAAACAAAGTGGCCATCCGCGCTTCGTC 61
80 TGGCTGGCCATGACATCACTTACACAGCAAG-GGCTTAATCCGCTTACGCTTC 138
62 TCGTCCCTCTCTGCTGCTGGGAGGAGCATCTACAGCTGCTGTATGAGAAATTCCT 121
139 TCCCGCTGCTGCTGCTGGGAGGAGCATCTACAGCTGCTATATGCGAGTTCTTA 198
122 GTCCTATATTCCTCTACTATTCATCCGAGCATCTACAGATGTTCTTCGAGTAT 181
199 ATCTTCTGCTGCTGCTACTACATCCGCTTATTTATGAGCTGCTCCCTACGGAAGA 258
182 CAGCAGCTGTTTGTGAGAGCTGCTCTGTACTGAGCAGCTACATTCAGTCCCT 241
259 CACAGCTGATGTTTGAAGAACTGACTGTGATGAGACAGCTACATCCAGTCTCC 318
242 ATATCCTGCTGCTGGGTTTCTATATGTAATGTTGTTGAGCGCTGTTGAGCCATAC 301
319 ATTTCTTGTGCTGGGCTTCTACGTAAGCTGCTGCTGAGCCGCTGTTGGAACAGTAC 378
302 GAGAACTTCCGCTGGCCGAGCCGCTCATGATCCAGTGTCTAGCTTGTGGAGGGCAG 361
379 GAGAACTTCCGCTGGCCGAGCCGCTCATGAGCTGCTGCTGAGGCTTGTGAGAGCAG 438
362 GATGAGGAAGGCGCTTGTGCTGGGAGCAGCTCATCCGCTACGCCCTTGTGGCCAGT 421
439 GACGAGCAAGGCGGCTGCTGCGGAGCAGCTCATCCGCTACGCCAATCTGGGAGGCTG 498
422 CTGATCTGCGCAGATGAGCAGCTCGGCTTACAGAGGCTTCCACCTTTCACACCTG 481
499 CTGATCTGCGCAGATGAGCAGCTCGGCTTACAGAGGCTTCCACCTTTCACACCTG 558
482 GTGCTAGAGGTTTATGAGCCATGAGGAGCAATAGCAGTGTGAGGCTTACCA 541
559 GTGCAAGCAGGCTTATGACTCCGAGCAGACAGAGGAGTTGGAGAACTGAGCTTACA 618
542 CACAAACATTTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
619 CACAAACATTTCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 678
602 GAGAGTCGAATCCGAGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
679 GAGAGTCGAATCCGAGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 738
662 GCTACTAGTGTGAGCAGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721
739 GCTACTAGTGTGAGCAGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
722 CAGTGTGAGCAGTGTGAGTATACAGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781
799 CAGTGTGAGCAGTGTGAGTATACAGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858
782 CTGAAACCAACAGAGCTACCCAGGAGCATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 841
859 CTGAAACCAACAGAGCTACCCAGGAGCATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 918
842 ATCTGAGTATTTATTTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
919 TTCTGAGTATTTATTTATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 978

QY 902 TTCCGGGAGAGCATGATGATTTTGAAGTAACTAGATCATTTGACAGAAACCTGACAGT 961
DB 979 TTGAGAGAGATGATGATTTTGAAGTAACTAGATCATTTGACAGAAACCTGACAGT 1038
QY 962 TTCCCTGTTGCTCCGTTGATGAGATGACACAACTTGGCTCCCATGAGAACTGACATGAT 1021
DB 1039 TTCCCTGTTGCTCCGTTGATGAGATGACACAACTTGGCTCCCATGAGAACTGACATGAT 1098
QY 1022 TGAAGAGAGGAGGCTCCAGCCGCTTACACAGCTGCTTGTGCAAGTGTGCTGCTGCTGCT 1081
DB 1099 TGAATTAAGCCCGAGGAGCAGACCCCTTACAGAGCTGCTTCCGCTGCTGCTGCTGCT 1158
QY 1082 TCCCTTATGAGGCTCCAGCTTCAATCATGAGCTTAAAGAAAGACTTAAAGCTTGTGCTA 1141
DB 1159 TCCCTTATGAGGCTCCAGCTTCAATCATGAGCTTAAAGAAAGACTTAAAGCTTGTGCTA 1218
QY 1142 AAAGAGAGGCTGACAGGATTAAGAAAGAGTGGCTATGAGCAGCAGCATAGGCTGCTT 1201
DB 1219 AATCAGAGAGCAGAGAGATGCTCAGCTG-----GCATCATTTGCTGCTGCTT 1266
QY 1202 TTAGAGCTGCAACCCCAAAAATACCATCTTCCCTTGAAGACTTAAAGCAAACTATTG 1261
DB 1267 CTAGGCTGCAAGTCCCATGATCAACATCTCCAGGAGCAAACTCAAGACCAATCTACTG 1326
QY 1262 TGTCTTAAAG-----ACCCCTCTGGAAGGCTGTAAGATGACCAAGAAAC 1315
DB 1327 TGGCCCAAGAGGAGATCCCTTCTCAGAGAGGCTGCTCCCAAAACCAAGGAGCCAA 1386
QY 1316 CAGAAAG-----ATGCTGGAATTTAAGGCTGAGCTTCTG 1354
DB 1387 CAGAGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1446
QY 1355 AAATGTTTCAAGGTTTAAAGAGAGAGGCTCCCATTTGAGGCTGAGGAGGAGGAGG 1414
DB 1447 AAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1506
QY 1415 CACCTTAC 1422
DB 1507 CCACTTCC 1514

RESULT 11
BC041664
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BC041664
Homo sapiens vitelliform macular dystrophy (Best disease, bestrophin), mRNA (cDNA clone MGC:47884 IMAGE:5194649), complete cds.
BC041664
BC041664.1 GI:27371319
MGC.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2435)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Waki, S.I., Wang, J., Hsieh, F., Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Tishiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wortley, K.C., Hale, S., Garcia, A.M., Gay, L.U., Hu, Y., S.W., Villalón, D.K., Wuzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Buterfield, Y.S., Krzywinski, M.I., Skalska, U., Smolins, D.E.,

TITLE Schnerch, A., Schein, J. E., Jones, S. J., and Marra, M. A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL MEDLINE
22388257
12477932
2 (bases 1 to 2435)
Straube, R.
Direct Submission
Submitted (20-DEC-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nih.gov
Alster, N., Ayala, K., Beckstrom-Sternberg, S. M., Benjamin, B.,
Blakesley, R. W., Bouffard, G. G., Breen, K., Brinkley, C., Brooke, S.,
Dietrich, N. L., Granite, S., Guan, J., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, F., Legaspi, R.,
Maduro, Q. L., Masiello, C., Maekel, B., Mastrian, S. D., McCloskey, J. C.,
McDowell, J., Pearson, R., Stantipod, S., Thomas, P. J., Touchman, J. W.,
Taurgeon, C., Vogt, J. L., Walker, M. A., Wecherby, K. D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E. D.

FEATURES
source
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 82 Row: d Column: 4.
Location/Qualifiers
1..2435
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:47884 IMAGE:5194649"
/tissue_type="Brain, adult, 6 pooled whole brains"
/clone_lib="NIH MGC_114"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2435
/gene="VMD2"
/note="synonyms: BMD, BEST"
/db_xref="LOCUSID:7439"
/db_xref="MIM:153700"
87..1901
/gene="VMD2"
/codon_start=1
/product="VMD2 protein"
/protein_id="AAH41664.1"
/db_xref="GI:27371320"
/db_xref="LOCUSID:7439"
/db_xref="MIM:153700"
/translation="MFEKLTLYCDSYQLIPISFVLGRVYLVVTRWNOYENLPMWP
RLMSLVGFVGEKDEGRILRLTLRYANLVNLSVSTAVVKKFSPSAHLVQAF
MTPEAHKOLEKLSLPHNMFWPMVPMFANLSKAWLGGIRDP2ILLOSILNMTLRG
CGHLYAVDWISLPLVYQVYVAVYSPFLTGLGROPRLNPKAKGHEHLLVVPFTE
LQPFYVGLKVAEQLINPGEDEDDDEFTNIVPRNLQVSLAVDENKODLPRNEPDM
YWNKRPQPTAAQFRRVRSFMSFTNISLNEKEMEPQNGDESDAAAGILGRFL
GLQGHDPHPRANSTRKLLMPKRESLHEGLPKXHKAKAKVNRQDEBNKAAVAVK
KSPALYQRPYUAPOTPLSPFMPFLBPSAKLSVYGIIDKDKSLTVSGAK
KSFELSESDALMEHPESQVRKRVTFENLTDMETPEBNILKEPLBESPTNHTTK
DHMDPVALNRSVLIHNOGCIACPTPALSIALISPLINHLFGLHHQCSQSLDLRAL
AKGILATFTGILKCSGPFUTSPWYHREDLGRGGR"
90..1022
/gene="VMD2"
/note="DUF289; Region: Putative membrane protein. This

family called family 8 in, may be a transmembrane protein
The specific function of this protein is unknown"
/db_xref="CDD:pfam01062"

ORIGIN

Query Match 40.6%; Score 777.2; DB 9; Length 2435;
Best Local Similarity 76.9%; Pred. No. 1e-178;
Matches 1004; Conservative 0; Mismatches 263; Indels 39; Gaps 3;

144 CCATCCGTGACCTCTACAGATGGTTCTCGAGTATGATCAGAGCTGTTGTTGAAGC 203
|||
40 CCCACCTGGAACTCCACCTGCTGGCCCTCAGAGAAACACAGCTGATGTTGAAGC 99
|||
204 TGGCTGTGACTGGACAGCTACATTCAGCTCATCTCTATATCTTCCTTGGGTTCT 263
|||
100 TGAATCTGTATGAGACAGCTACATTCAGCTCATCTCTATTCCTTGGTGGCTTCT 159
|||
264 ATGTTACATGGTGTGAGACCGCTGTGTGAGACCCATGACAGAACTTGCGTGGCCGAC 323
|||
160 ACGTAGCGTGTGTGAGACCGCTGTGTGAGACCCATGACAGAACTTGCGTGGCCGAC 219
|||
324 GCTTCATGATCAGAGTGTGTGAGTGTGTGAGAGGCAAGATGAGGAAGCCGTTTGGCTG 383
|||
220 GCTTCATGAGCTGTGTGTGAGTGTGTGAGAGGCAAGATGAGGAAGCCGCTTGC 279
|||
384 GGCACAGCTCATCGCTACAGCTACCTCTGAGGCAAGTCTCATCTGCGACATACAGA 443
|||
280 GGGCCACGCTCATCGCTACAGCTACCTCTGAGGCAAGTCTCATCTGCGACGCTGAGCA 339
|||
444 CTTGCTGTACAGAGCTTTTCCACTTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 503
|||
340 CCGAGCTTCAAGAGCTTTTCCACTTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399
|||
504 ATGGGAAACATTAACAGTGTGAGAGTGTGTGAGAGGCAAGATGAGGAAGCCGTTTGGCTG 563
|||
400 CGGAGAAACATTAACAGTGTGAGAGTGTGTGAGAGGCAAGATGAGGAAGCCGTTTGGCTG 459
|||
564 GGGTGTGTGTGAGAGCTTTTCCACTTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 623
|||
460 GGGTGTGTGTGAGAGCTTTTCCACTTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519
|||
624 TCTGCTTCAGAGCTGTGAGAGTGTGTGAGAGGCAAGATGAGGAAGCCGTTTGGCTG 683
|||
520 TCTGCTTCAGAGCTGTGAGAGTGTGTGAGAGGCAAGATGAGGAAGCCGTTTGGCTG 579
|||
684 ATGCTTACAGTGTGAGAGTGTGTGAGAGGCAAGATGAGGAAGCCGTTTGGCTG 743
|||
580 ATGCTTACAGTGTGAGAGTGTGTGAGAGGCAAGATGAGGAAGCCGTTTGGCTG 639
|||
744 ACAGCTTTTCTGTCATGCTGTGAGAGGCAAGATGAGGAAGCCGTTTGGCTG 803
|||
640 ACAGCTTTTCTGTCATGCTGTGAGAGGCAAGATGAGGAAGCCGTTTGGCTG 699
|||
804 CAGGCAATGAGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 863
|||
700 CTGGCAATGAGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 759
|||
864 TGGCTGTGAGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 923
|||
760 TGGCTGTGAGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 819
|||
924 TTTGAGCAATGAGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 983
|||
820 TTTGAGCAATGAGATGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 879
|||
984 TGCACAGAACTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1043
|||
880 TGCACAGAACTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 939
|||
1044 CGCCTTACAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1103
|||
940 CCCCTTACAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 999
|||

QY	1104	ACATCAGCCTAAAGAAAGAAAGCTTAGAGCTTTGGTCAAAGAGAGAGCGCTACAGGATA	1163
Db	1000	ACATTAGCCTGAACAAAGAGGAGATGAGATTCCAGCCCAATCCAGAGG-----	1047
QY	1164	AGAAAGAGATGGCTATAGACGACACATAGGCGTGTCTTTAGGATGTGCAACCCAAAAC	1222
Db	1048	ACGAGGAAGATGCTACGCTGGCATCATTTGGCGGCTCTTAGCGCTGACATGCCATGATC	1107
QY	1224	ACCATCTTCCCTTGGAAAGACTTTAAAGACCAAACTATTTGTGTTCTAAG-----ACCC	1277
Db	1108	ACCATCTCTCCAGGGCAAACTCAGAGACCAAACTACTGTGGCCCAAGGAGAAATCCTTC	1167
QY	1278	TCCTGGAAGCCAGTGAAGATGCGCAACAGAAAAACAGAAAG-----	1322
Db	1168	TCCAGAGAGGCGTGGCCAAAAACCAAGAGGACGCAACAGAGCTTAGGGGCCAGAG	1227
QY	1333	-----ATGCTCGAAATTTAAGGCTCTGCACTTTGAAATGTTGTTCCAAAGTTTAAG	1376
Db	1228	ACAAACAAAGGCGTGAAGCTTTAAGGCTGTGGACGCTTCAAGCTCTGCCCACTGATACAG	1287
QY	1377	GGAGAGGCTCCATTTGTGGCCCAACGGACCCAGACGCCACTTAC	1422
Db	1288	GGCACGAGCTACTACAGTCCCAACAGAGGCCCTCTAGCCCACTCC	1333

RESULT	12
LOCUS	AF057170
DEFINITION	Homo sapiens bestrophin (VMD2) mRNA, alternatively spliced product,
ACCESSION	AF057170
VERSION	AF057170.1
KEYWORDS	GI:3335160
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Petrushkin,K., Koistinen,M.J., Bakall,B., Li,W., Xie,G., Martnell,T., Sandgren,O., Forsman,K., Holmgren,G., Andersson,S., Vujic,M., Bergen,A.A., McCarley-Dugan,V., Figueroa,D., Austin,C.P., Metzker,M.L., Caskey,C.T. and Wadelius,C. Identification of the gene responsible for Best macular dystrophy Nat. Genet. 19 (3), 241-247 (1998) 98324772
JOURNAL MEDLINE	9662395
PIUMED	2 (bases 1 to 2420)
REFERENCE	Petrushkin,K. Direct Submission Submitted (03-APR-1998) Human Genetics, Merck Research Laboratories, West Point, PA 19486, USA
AUTHORS TITLE JOURNAL	Location/Qualifiers
FEATURES	1..2420
source	/organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="11" /map="11q13" 1..2420 /gene="VMD2" 105..1412 /gene="VMD2" /note="Best's macular dystrophy gene; BMD; alternatively splice isoform; contains exon 7"
gene CDS	/codon_start=1 /product="bestrophin" /protein_id="AAC64344.1" /db_xref="GI:3335161" /translation="MTITTVSYOVNARLGSFRLLCWRGSIYKLIGEPILFLCY IIRPLALALEGOQLMEKTLICDSEYLIOLIPISFYLGYYVLIVTRMNNOVENLPW GPPTLSMSVSFGVEKDRQSGULRRLLRIYRNALGNVILRSVSTRVYKRFSAGHLVOA PGRTPAHKKOLEKSLDFHNHFVPWFVAFLSMKAWIGSRIDPTLLOSILNENNTLR TGGCHLAAYMOEISLPIVYTGVTVAVYSFFLTCAVGQFLNPAKAYGHELDLVPEVF

ORIGIN

Query Match	34.4%;	Score 658.2;	DB 9;	Length 2420;
Best Local Similarity	67.6%;	Pred. No. 1.3e-149;		
Matches 116;	Conservative 0;	Mismatches 293;	Indels 242;	Gaps 4;

TELQFFFYVWGLKVLGRALGWRHGORGQOQLLETRMOCOERKYSRVESSOAMWRT
PVIPATREAAGESLEPGRRLWMOSSSTPLRMMMLIRPTGLSTGICRCPCLMMR
CTRCLGWSRTCTGISPSHSPPTQLLPSSVEPPLMAPDPSTA "

QY	2	GGCCAGGCAATGACTATACCTTACATACAAACAAGTAGGCATATGCCCGCTGGTTCC	61
Db	96	TGCCGCGCAATGACATCATCTTAAACAAGCAAGTGGTAAATGCCCGCTTATGGCTCTTC	155
QY	62	TGCTCCCTCCCTCCGTGTGCGAGGCGAGCATCTACAAGCTGCTGTATGAGAAATTCCTT	121
Db	156	TCCCGCCTGCTGTGTGTGTGTGGCGGGGAGCATCTACAAAGCTGTATATGGGAATTCCTA	215
QY	122	GTCTTCATATTTCTCTACTATTCATCCGNGAACTTACAGATATGTTCTCTCGAGTAT	181
Db	216	ATCTTCCTGCTGTGCTACTACTATCATCATCCGCTTTATTTATATAGCTGGGCCCTCACGGAA	275
QY	182	CAGACAGCTGTGTGTTGAGAAAGCTGGCTGTATCTGCGACAGCTACATTTCACTCATCCCT	241
Db	276	CAACAGCTGATGTTTGGAAACTACACTCTGATTTATTCGACAGCTACATCCACTATCCCC	335
QY	242	ATATCCCTTCGTTCTGGGTTTCTATAGTTACATTTGGGNGAGCCGCTGGTGGAGCCAGTAC	301
Db	336	ATTTCCTTCGTGTGCTGGCTTCTTACGTACGCTGTGTCTGACCCGCTGTGGAAACAGTAC	395
QY	302	GAGAACCTTGCCTGGGCCGACCCGCTCATATGATCCAGGTGTCTAGCTTTGTGGAGGGCAAG	361
Db	396	GAGAACCTGCGGTGGCCCGACCCGCTCATATGAGCTGTGGTGTGGGCTTGTGGAAGGCAAG	455
QY	362	GATATGAGAAAGGCCCTTTGTCTGGCGGCGACAGCTCATCCGCTACCGCATCTCTGGGCGCAAGT	421
Db	456	GATCAGAGCAAGCCGCGCTGTGCGGGCGGACGCTCATCCGCTACCGCAACTGGGGCAACCTGT	515
QY	422	CTCATCTCCGCGACATACAGCATCTCGAGTCTACAAGGCTTTTCCCATCTTCAACAACCTGT	481
Db	516	CTCATCTCTGCGACGCTACAGCACCGGACGTACAAAGGCTTCCCAAGGCCACGACCTGT	575
QY	482	GTGCTAGCAGGTTTATATGACCCATGGGGAACATTAAGCAGTTGAGAAATGGGCTTACCA	541
Db	576	GTGCAACAGGCTTTATATGACTCCGGGAGAAACAACAAGCAGTTGGAGAAACTGAGCCTACCA	635
QY	542	CACAAKCAATTTGGGGTGCCCTGGGTGTGTGTTTCCCACTTGTCAATGAAAGCCTATATT	601
Db	636	CACAAKCAATTTGGGGTGCCCTGGGTGTGTGTTTCCCAACTGTCAATGAAAGCCTGTGGCTT	695
QY	602	GGAGGTGGAATCCGGGAGCACCGTCTGTCTCCAGAGCCGTGATGATGAGGTGTACTTTG	661
Db	696	GGAGGTGGAATCCGGGAGCCCTTATCTGTCTCCAGAGCCTGTCTGAAAGATGAAACAACCTTG	755
QY	662	CGTACTCAGTGTGAGACAGCTGTATATGCTTACGACTGTATTAAGTATCCCATTTGTTACACA	721
Db	756	CGTACTCAGTGTGAGACACTGTATATGCTTACGACTGTATTAAGTATCCCATGTTGTATAC	815
QY	722	CAGGTGTGTACAGTGGCAGTATACAGCTTTTCTTTCGATATGCTTGAATGGGAGGCAAGTTT	781
Db	816	CAGGTGTGTACAGTGGGAGTGTACAGCTTTCTTCTGTACTGTCTAGTGTGGCGGCAAGTTT	875
QY	782	CTGAATCCCAACAAGACTACCCAGAGCCATGAGATGGAATCGTGTGGCCGTGTTTACA	841
Db	876	CTGAATCCCAACAAGGCTTACCCCTGGGCGCATGAGTGAAGCTCGTTGTGCCCGTTCACG	935
QY	842	ATCCTGCAATTTCTTATTTCTACATGAGGCTGAGCTGAA	876
Db	936	TTCCTGTGAGTTCTTCTTCTATATGTTGGTGTGCTGAAGTGGGCCCTCTCACAGGCCCTGTG	995
QY	877	-----	876
Db	996	GGCTGGAGCATGAGCCAGAGGGGTCAATGAGCCAGACGCTGTGTTGAGACGAGATGCAAGT	1055

877 ----- 876
Qy 877 -----
Db 1056 CAGGAAAGAGGTCTCA CGGGTAGAAAAGACAGCGCGTGTGGCGCACCTGTAATC 1115
Qy 877 -----GG 878
Db 1116 CCAGCTACTCGGAGGCTGAGGACAGAGAAATCGCTGAACCCGGAGCGGAGGTTGTGG 1175
Qy 879 TGGCAAGACAGCTATCAACCCCTTCGGGAGAGACGATGATGATTTTGAAGTAATCTGGA 938
Db 1176 TGGCAGAGCAGCTATCAACCCCTTCGGGAGAGATGATGATTTTGAAGTAATCTGGA 1235
Qy 939 TCATTTGACAGAAACCTGAGGTGTCCCTGTTCCTGTGATGGAATGCAACCAAGATTGC 998
Db 1236 TTGTGACAGGAAATTTGACAGGTGTCCCTGTTCCTGTGATGATGATGACCAAGGCTGC 1295
Qy 999 CTCCATGGAACGTGACATGTACTGGAACGAGGAGCGCTCAGCCGCTTACACAGTGT 1058
Db 1296 CTCGATGAGACCGGACATGTACTGGAATTAAGCCGAGCCACAGCCCTTACACAGCTG 1355
Qy 1059 CTTCGACAGGTCTGCGCGGATTCCTTCATGGGCTCCAGCTCAACATCAGCCTAAAGA 1118
Db 1356 CTTCGCGCCAGTTCCTTCGAGGCTTCCTTTATGGGCTCCAGCTCAACATCAGCCTGAACA 1415
Qy 1119 AAGAAAGCTTAGAGCTTTGTCAAAAAGAGAGGCTGACACGAGATTAAGAAAGAGTGGCT 1178
Db 1416 AAGAGAGATGAGATTCAGGCCAATCAGAGG-----ACGAGAGAGATGCTC 1463
Qy 1179 ATAGACAGCATTAGGCTGCTTTTGAAGATGCAACCAAAACTACCATTTTCCCTTGA 1238
Db 1464 ACGGTGATCATTTGGCGGCTTCCTAGGCTGACAGTCCCATGATCACCATCTCCCAAGG 1523
Qy 1239 AAGATTAAGACCAAACTATTTGTCTTAGAA-----CCGCTCTCGAAGGCGCAGT 1292
Db 1524 CAAACTCAAGACCAAACTATGTCGCCCAAGAGGAAATCCCTTCTCCACAGAGGCTGC 1583
Qy 1293 GTAAGATGCCAACAGAAAAACAGAAAGAT-----GTCTGGA 1331
Db 1584 CCAAAACCAAGAGCAGCCAAACAGAAAGTTAGGGGCAAGAAAGACAAAGGCTTGA 1643
Qy 1332 AATTTAAGGCTGTGACTTTTGAATGTGTTCCAGGTTTAAAGAGAGGCTCCATT 1391
Db 1644 AGCTTAAGGCTGTGAGCGCTTCAAGTCTGCCCACTGTATCAGAGGCGAGCTACTACA 1703
Qy 1392 GTGGCCACAGGCAAGCCAGAGCCACCTAC 1422
Db 1704 GTGCCCAACAGCGCCCTCAGCCCACTCC 1734

RESULT 13
BD136721
LOCUS BD136721 2429 bp DNA linear PAT 18-SEP-2002
DEFINITION Best's macular dystrophy gene.
ACCESSION BD136721
VERSION BD136721.1 GI:23231666
KEYWORDS JP 2002504559-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2429)
Petrushkin, K., Caskey, T.C., Metzker, M. and Madeline, C.
Best's macular dystrophy gene
Patent: JP 2002504559-A 3 12-FEB-2002;
MERCK & CO INC, CLAS WADLIUS
OS Homo sapiens (human)
PN JP 2002504559-A/3
PD 12-FEB-2002
PR 25-FEB-1998 JP 2000533447
PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUSHKIN, THOMAS C CASKEY, MICHAEL METZKER, CLAES PI
WADLIUS
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/68// PC

C12P21/08,
PC C12N5/00, C12N15/00
CC Best's macular dystrophy gene
FH Key Location/Qualifiers
FT source 1..2429 /organism='Homo sapiens (human)',
FEATURES
source location/Qualifiers
1..2429
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 34.4%; Score 658.2; DB 6; Length 2429;
Best Local Similarity 67.6%; Pred. No. 1.3e-149;
Matches 116; Conservative 0; Mismatches 293; Indels 242; Gaps 4;
Qy 2 TGGCAAGCATGACTATACCTTACCAAAACAAAGTAGCCATGCCCCGCTGCTTCGTTTC 61
Db 96 TGCCTGGCATGACCATCACTTACAAAGCCAAAGTGGCTAATGCCCTTAGGCTCCCTTC 155
Qy 62 TCGTCCCTCCCTGCTGCTGCTGGGAGAGCATTTAAAGCTGTATATGAGAAATTCCTT 121
Db 156 TCCGCTGCTGCTGCTGCTGGGAGAGCATTTAAAGCTGTATATGAGAAATTCCTT 215
Qy 122 GTCTCATATTCCTCTACTATATTCATCCGAGACTTACAGAAATGTTCTTCGAGTAT 181
Db 216 ATCTTCTGCTGCTGCTACTATATTCATCCGCTTATTTATAGGCTGGCCCTACAGGA 275
Qy 182 CAGCAGCTGTTTGTGAAGAGCTGCTGTACTGCGACAGCTACATTCAGCTCATCCCT 241
Db 276 CAACAGCTGATGTTTGAAGAACTGACTGTGATGCGACAGCTACATTCAGCTCATCC 335
Qy 242 ATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 301
Db 336 ATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 395
Qy 302 GAGAACTTGGCCGCTGCGCCGAGCCGCTCATGATCCAGGCTGTACTGTTGAGAGGCG 361
Db 396 GAGAACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 455
Qy 362 GATGAGAAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
Db 456 GAGCAGAAAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 515
Qy 422 CTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
Db 516 CTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 575
Qy 482 GTGCTAGAGGTTTATAGCCATGAGGAGAAACATAGAGAGTTGAGAAAGTTGGGCTACCA 541
Db 576 GTGCAAGCAGGCTTTATAGCTCCGCGAGAACAGAGATTTGAGAAAGTTGAGGCTTACCA 635
Qy 542 CACAACATTTCTGGGATGCCCTGGGTGTGTTTGCACATTTGTCAATGAAAGCCTATCTT 601
Db 636 CACAACATTTCTGGGATGCCCTGGGTGTGTTTGCACATTTGTCAATGAAAGCCTATCTT 695
Qy 602 GAGAGTGAATCCGGACACCGTCTGCTCCAGAGCCTGATGAATGAGGTGTATCTTGG 661
Db 696 GAGAGTGAATCCGGACACCGTCTGCTCCAGAGCCTGCTGAGAGAGATGAAACCTTGG 755
Qy 662 GTATCTAGAGTGGACAGCTGATAGCTGATAGCAAGATGATAGTATCCCATTTGGTATACA 721
Db 756 GTATCTAGAGTGGACAGCTGATAGCTGATAGCAAGATGATAGTATCCCATTTGGTATACA 815
Qy 722 CAGGTGTGACAGTGGCAGTATACAGCTTTTCTTCAGTCTGATCGGAGGACAGTTT 781
Db 816 CAGGTGTGACAGTGGCAGTATACAGCTTTTCTTCAGTCTGATCGGAGGACAGTTT 875
Qy 782 CTGAACCAAAACAAAGACTACCAAGCCATGAGATGATGATGATGATGATGATGATGAT 841
Db 876 CTGAACCAAAACAAAGACTACCAAGCCATGAGATGATGATGATGATGATGATGATGATG 935


```

QY      842 ATCTGCAATTCCTATTCATCATGCGCTGGCTGAA----- 876
      ||||| ||||| ||||| ||||| |||||
Db      936 TTCTCTGAGTTCTTCTTCTATGTTGGCTGCTGAAGGTGGGCTCTCCAGGCGCCCTG 995
QY      877 ----- 876
Db      996 GGCTGAGCATGGCCAGAGGGGTCAATGCGCAGAGCTGTTGAGACGAGAGTCACTGT 1055
QY      877 ----- 876
Db      1056 CAGGAAGAAGGTCTCAAGGGTGAAGAACAGCAGCGGTGGCCGACACTGTATC 1115
QY      877 -----GG 878
Db      1116 CCAGCTACTGGGAGGCTGAGGAGCAAGAAATCGTTGAACCCGGAGGCGGAGTTGTGG 1175
QY      879 TGGCAGAACAGCTCATCAACCCCTTCGCGGAGAGCATGATGATTTTGAAGTACATGGA 938
Db      1176 TGGCAGAGCAGCTCATCAACCCCTTGGAGAGGATGATGATTTTGAAGCAACTGGA 1235
QY      939 TCATTTGACAGAAACCTGAGGTCTCCCTGTTCTCCGATGAGGATGACACAGAACTTGC 998
Db      1236 TTGTGACAGGAATTTGAGGTGTCCTGTTGGCTGTGATGAGATGACACAGAACCTGC 1295
QY      999 CTCCTCATGAAACGTGACATGTACTGAAACGAGCGAGCGCTCAGCCGCTTACACAGCTG 1058
Db      1296 CTGCGATGAGCGCGGACATGTACTGGAATAAGCCCGAGCCACAGCCCTTACACAGCTG 1355
QY      1059 CTTCGCGAGGTCTGCGCGGCAATTCCTTCATGSGCTCCACCTTCAACATCAGCTTAAAG 1118
Db      1356 CTTCGCGCGAGTTCTCGTGAAGCTCTCTTTATGSGCTCCACCTTCAACATCAGCTTGAACA 1415
QY      1119 AAGGAGCTTAGACCTTTGGTCANAAGAGAGGCTGACAGGATTAAGAAAGAGTGGCT 1178
Db      1416 AAGGAGGATTTGAGTTCCAGCCCATCAAGAGG-----ACGAGGAGGATGCTC 1463
QY      1179 ATAGACACACCATAGAGCTGCTTTCTTAAAGATGCAACCAAAATACATCTTCCCTTGA 1238
Db      1464 ACGTGGCATCATTTGGCGGCTTCTTAAGCCCTGAGTCCCATGATCAGCATCTTCCAGGG 1523
QY      1239 AAGACTTAAGACCAAACTATTTGTTCTTAAAG-----CCCCCTCTCGAAGGCCAGT 1292
Db      1524 CAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGGAATCCCTTCCACAGAGGCGCTGC 1583
QY      1293 GTAGAGTGCACACAGAAAACAGAGAAAGT-----GTCGTGA 1331
Db      1584 CCAAAAACCAAGGACGACCAAGAAAGCTTAAGGGCCAGAAAGACAAAGAGCCTGGA 1643
QY      1332 AATTAAAGGCTCTGACTCTTGAATGTGTTCCAGGTTTAAAGAGAGAGGCTCCATTT 1391
Db      1644 AGCTTAAGGCTGTGAGAGGCTTCAAGTCTGGCCCACTGTATCAAGAGGCGAGCTACTACA 1703
QY      1392 GTGGCCCAAGGACCAAGCAGAGCCACCTTAC 1422
Db      1704 GTGCCCAAGAGCGCCCTTCAAGCCCACTCC 1734

```

```

RESULT 14
BC015220      2441 bp      mRNA      linear      PRI 07-OCT-2003
LOCUS      BC015220
DEFINITION      Homo sapiens vitelliform macular dystrophy (Best disease,
bestrophin), mRNA (cDNA clone MGC:17761 IMAGE:3877806), complete
cde.
ACCESSION      BC015220
VERSION      BC015220.2
KEYWORDS      GI:34783515
SOURCE      MGC.
ORGANISM      Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

```

```

TITLE
JOURNAL
MEDLINE
PUBMED
2388257
12477932
2 (bases 1 to 2441)
Strausberg,R.
REFERENCE
AUTHORS
JOURNAL
TITLE
REMARK
COMMENT
NHL-MGC Project URL: http://mgc.nci.nih.gov
On Sep 16, 2003 this sequence version replaced gi:2195361.
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E Consortium (LINL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E Consortium/LINL at: http://image.lnl.gov
Series: IRAK Plate: 14 Row: b Column: 9.
Location/Qualifiers
1..2441
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:17761 IMAGE:3877806"
/tissue_type="Lung, large cell carcinoma"
/clone_lib="NIH MGC_68"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1..2441
/clone="VMD2"
/note="synonyms: BMD, BEST"
/db_xref="LOCUSID:7439"
/db_xref="MIM:153700"
573..2069
/clone="VMD2"
/codon_start=1
/product="VMD2 protein"
/protein_id="AAH15220.1"
/db_xref="GI:2280061"
/db_xref="LOCUSID:7439"
/db_xref="MIM:153700"
/translaction="MBEKLTYCDSYIOLIPISFVLGFTYTLVVTMMNOYENLPMPD
RMSLSVGFVKGDEGRLRLIRANGLVILRSVTAAYKRPASOHLVOAGF
MTFAEKOLEKLSLPHRMFVPMVPMRANLSMKWLGRIIDPILLOSILAMNTLRFO
CGHUVYDWISLVLVYQVTVAVISFELCLVGRPLPAKAYRGHEDLVVDFVTF
LQFFVYGMVLKSLAVDEHODLPKMEPDPMWKNKPPPPYTAASQFRASPMGST

```


FNLSLNKEMEFQNOEDBEDAHAGIIGRFLQSHDHPPRANSRTKLMPKRESLI
HEGLPKHKAQONVGOEDBNKAWLKAVDFKSAPIYORGYYSAQOTPLSPMPFE
PLEPSAPSKLSVTDIKDLSKTIVSSAKKSEPLISESGALMEHPVSQVARKTV
ERNITDMEPIBENHLKEPLESPTNIHTTLKDHPDYMALENRDEAS"
576..1427
/gene="VMD2"
/note="DUF289; Region: Putative membrane protein. This
family called family 8 in, may be a transmembrane protein
The specific function of this protein is unknown"
/db_xref="CDD:pfam01062"

ORIGIN

```
Query Match      32.6%; Score 624.4; DB 9; Length 2441;
Best Local Similarity 71.6%; Pred. No. 2.3e-141;
Matches 935; Conservative 0; Mismatches 251; Indels 120; Gaps 4;
```

144 CCATCCGTGAGCTTACAGAAATGTTCTCTGAGTATCAGCAGCTGTTGTTGAAGC 203
Db CCCACCTGGAACCCACCTGCTGCCCCCTCAGGAAGAACACCTGATGTTGAAGAAC 585
Qy TGGCTCTGTAATGGAGACGCTACATTCAGCTATCCCTATCCCTGCTGCTGCTTCT 263
Db TGAATCTGTATTTGGACAGCTACATCAGCTATCCCTATCCCTGCTGCTGCTTCT 645
Qy 264 ATGTTAATTTGGTGTGAGCCGCTGTGTGAGCAGTACAGAACTTGGCCGTGAGC 323
Db 646 ACGTGAAGCTGTGTGAGCCGCTGTGTGAGCAGTACAGAACTTGGCCGTGAGC 705
Qy 324 GCTCATATGATCAGAGTGTCTAGCTTCTGTGAGGGCAAGATGAGGAAGCCGTTTGTCTG 383
Db 706 GCTCATATGAGCTGTGTGTGAGGGCAAGATGAGGAAGCCGCTGCTGCTG 765
Qy 384 GGGCAGCTCATCCGCTACAGCCATCCGAGCCCAAGCTCATCTGTGGCAGCATCAGCA 443
Db 766 GGGCAGCTCATCCGCTACAGCCATCCGAGCCCAAGCTCATCTGTGGCAGCATCAGCA 825
Qy 444 CCTGGCTCTACAAAGCGCTTCCCACTCTTACCAAGCTGTGTGAGAGTTTATGACCC 503
Db 826 CCGCAGCTTACAAAGCGCTTCCCAAGCGCCAGCACTGTGTGAGAGCTTTATGATC 885
Qy 504 ATGAGGGAACATAGCAGTTGAGAAAGTTGGGCTTACACACAGACATCTTGGGTGCCCT 563
Db 886 CGGAGAGAACATAGCAGTTGAGAAAGTTGAGCTTACACACATGTTTGGGTGCCCT 945
Qy 564 GGGTGTGTGTTGGCCAACTTGTCAATGAAAGGCTTATCTTGGAGTGTGAAATCCGGGACACCG 623
Db 946 GGGTGTGTGTTGGCCAACTTGTCAATGAAAGGCTTGTGAGTGTGAAATCCGGGACACCTTA 1005
Qy 624 TCTGCTCCAGAGCGCTGATGAAATGAGGTGTATCTTGGCCGACTCAGTGTGAGCAGCTGT 683
Db 1006 TCTGCTCCAGAGCGCTGATGAAATGAGGTGTATCTTGGCCGACTCAGTGTGAGCAGCTGT 1065
Qy 684 ATGCTTACGACTGTGATAGTATCCATTTGATGTATACACAGAGTGTGACAGTGTAT 743
Db 1066 ATGCTTACGACTGTGATAGTATCCCATGTTATATACACAGAGTGTGACAGTGTAT 1125
Qy 744 ACAGCTTTTCTGTCATGCTTGTGATCGGAGGAGCTTCTGAACCCAAACAGAGCTATCC 803
Db 1126 ACAGCTTTTCTGTCATGCTTGTGATCGGAGGAGCTTCTGAACCCAAACAGAGCTATCC 1185
Qy 804 CAGGCACTGAGATGATGATGTTGCTGTCTTCACAATCTGCAATTTTATTTTACA 863
Db 1186 CTGAGCACTGAGATGATGATGTTGCTGTCTTCACAATCTGCAATTTTATTTTACA 1245
Qy 864 TGGGCTGGCTGAGAGTGTGAGAGACAGTCAACCCCTTGGGAGAGAGATGATGATTT 923
Db 1246 TGGGCTGGCTGAGAGTGTGAGAGACAGTCAACCCCTTGGGAGAGAGATGATGATTT 1257
Qy 924 TTGAGACTAATGATGATTTGACAGAAACCTGACAGGTGTCCCTGTTTCCGTGATGAGGA 983
Db 1258 -----AGGTGTCCCTGTTGCTGTGATGATGAGA 1284
Qy 984 TGACACGAACTTGCTCCATGAGACGTGACATGTACTGGAACGAGGACGCGCTGAC 1043

1285 TGACACGAGACCTGCTCCGATGAGACCGGACATGTAGTAATGAGCCGACAGC 1344
Qy 1044 GCGCCTTACACAGCTGCTTCTGACAGTGTGCTGCGGCACTCTTATGAGGCTTCACTTCA 1103
Db 1345 CCGCCTTACACAGCTGCTTCTGACAGTGTGCTGCGGCACTCTTATGAGGCTTCACTTCA 1404
Qy 1104 ACATCAGCTTAAAGAAAGACTTTAGAGCTTTGATGATGATGATGATGATGATGATGATGAT 1163
Db 1405 ACATCAGCTTAAAGAAAGACTTTAGAGCTTTGATGATGATGATGATGATGATGATGATGAT 1452
Qy 1164 AGAAGAGAGTGTGATGAGCAGCAGTATGAGCTTCTTGAAGCTGACCAAGAACT 1223
Db 1453 ACGAGAGAGATGCTACAGCTGAGCAGTATGAGCTTCTGAGCTGAGCTGAGCTGAGCTGAGCT 1512
Qy 1224 ACCATCTTCCCTTAAAGAAAGCTTTAAGACCAAACTTATGTTGTTCAAG-----ACCCCC 1277
Db 1513 ACCATCTTCCCTTAAAGAAAGCTTTAAGACCAAACTTATGTTGTTCAAG-----ACCCCC 1572
Qy 1278 TCTCTGAAGGCACTGTATGAGTGCACCAAGCAAGAAACAGAAAG-----1322
Db 1573 TCCAGAGGCTTCCCAAGAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1632
Qy 1323 -----ATGCTGGAATTTAAGGCTGTGACTTCTTGAATGTTTCAAGGTTTAAAGA 1376
Db 1633 ACAACAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 1692
Qy 1377 GGAAGGCTCCCATTTGTGCTGACAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 1422
Db 1693 GGCAGGCTTACTACAGTGCCTCCACAGACACCTTCAAGCCCACTTC 1738

RESULT 15
AX745966 2137 bp DNA linear PAT 14-MAY-2003
LOCUS AX745966
DEFINITION Sequence 3 from Patent WO03030922.
ACCESSION AX745966
VERSION AX745966.1 GI:30724620
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
Steuernagel, A., Bruemer, G., Fritsch, R., Eulenberg, K. and
Closek, T.
Beetrophin and beetrophin homologous proteins involved in the
regulation of energy homeostasis
Patent: WO 03030922-A 3 17-APR-2003;
Develogen Aktiengesellschaft fuer Entwicklungsbioologische Forschung
(DE)
FEATURES
source
1..2137
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

```
Query Match      27.1%; Score 519.4; DB 6; Length 2137;
Best Local Similarity 66.9%; Pred. No. 1.1e-115;
Matches 755; Conservative 0; Mismatches 371; Indels 3; Gaps 1;
```

9 CCATGATATACCTTACAAACAAAGTAGCAATGCCCGCTGGCTTCTGCTGCTCC 68
Db 210 CGATGACCGTCACTTACACAGCCCGAGTGGGAAAGCCCGCTTGGTGGCTTCCAGC 269
Qy 69 TCTCTGTGCTGGCGAGGAGCATCTTACAAAGCTGTGATGAGAAATCTTGTCTTCA 128
Db 270 TGTCTGCTACTGTGCGGAGGAGCATCTTACAAATCTTGTGCGAGAGCTGCTTCTCC 329
Qy 129 TATTCTCTACTATTTTCATCCGTGAGCTTACAGAAATGTTTCTGAGATGATCAGAC 188
Db 330 TTGGGTTCTATCATATGAGCGCTGAGTGTGCTGCTTGTGCTGACCGAAGGAGCAAGAC 389

```
QY 189 TGTGTTTGGAGAGCTGGCTCTGTACTGCGACAGCTACATTCAAGCTGATCCCTATATCCCT 248
Db 390 GCTACTTCGAGAGAGCTGTATGATTTATGTATACAGATATGCCAGCTCATCCCTGTCTCT 449
QY 249 TCGTTCTGGGTTTCTATGTTACATTTGGTGGAGCCGCTGGTGGAGCCAGTACGAGAACT 308
Db 450 TCGTGTCTGGCTTTTATGTACGCTGTGTGTGAACCGCTGTGTGAGCCAGTACCTATATGA 509
QY 309 TGCCTGTGCGCCGACCGGCTTATGATTCAGGTGTCTAGCTTGTGTGAGAGGCAAGATGAGG 368
Db 510 TGCCTGTGCGCCGACCGGCTTATGATGTGTGTGTGAGAGGCAAGATGAGGCAAGATGAGG 569
QY 369 AAGGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 428
Db 570 GCGGCGCGCTTACCGGCGCACTATGCGTACGAGGAGGCTGTGCGCGGTGTGTGTGTGTGT 629
QY 429 TGCAGAGATCAGCAGCTCGGTCTACAGGCGCTTCCCACTCTTCAACACCTGTGTGTGTGT 488
Db 630 TGCCTGTGCGCCGACCGGCTTATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 689
QY 489 CAGGTTTTATGACCCATGAGGAGACATAGCAGTTGCAAGTTGGGCTTACACACACACA 548
Db 690 CTGGGTTTATGACCCGAGAGAGGCGCAAGATTGAAACCTGAACTCATCTTACACACA 749
QY 549 CATCTGTGGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 608
Db 750 AGTACTGGGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 809
QY 609 GAATCCGGGAGACCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 668
Db 810 GCATCCCGGAGACAGGCGCTTAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 869
QY 669 AGTGTGACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 728
Db 870 AATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 929
QY 729 TGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 788
Db 930 TGACCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 989
QY 789 CAAACAGAGACTACCCAGGCGCATGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 848
Db 990 CGGCTCAGGGTTTAAAGAGACACAGCATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1049
QY 849 AATTCTTATTTCTACATGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 908
Db 1050 AGTCTTCTCTACCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1109
QY 909 AGGACGATGATGATTTTGAAGCTAACTGATGATGATGATGATGATGATGATGATGATGATG 968
Db 1110 AGGACGATGATGATTTTGAAGCTAACTGATGATGATGATGATGATGATGATGATGATGATG 1169
QY 969 TGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1028
Db 1170 TGGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1229
QY 1029 AAGCAGCGCTTACGCGCGCTTACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1085
Db 1230 CAGCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1289
QY 1086 TCATGGGCTTCACTTCAACATCAAGCTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCT 1134
Db 1290 TCCAGGGCTTCACTTTCATCACTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1338
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2005, 22:43:02 ; Search time 6161.57 Seconds

(without alignments)
11836.456 Million cell updates/sec

Title: US-09-622-964a-28

Perfect score: 1916
Sequence: 1 gtgccaaagcatgactatca.....aaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_esc1:*
2: gb_esc2:*
3: gb_hic:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_g981:*
9: gb_g982:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1684.8	87.9	1699	3	AK006549
2	573.6	29.9	963	5	B0879880
3	519	27.1	713	6	B714331
4	502.6	26.2	881	6	CD518675
5	493.6	25.8	792	7	CK470840
6	488.2	25.4	756	7	CO396179
7	487	25.4	664	6	BY714344
8	474.2	24.7	862	7	CK469826
9	466	24.3	578	1	AU040308
10	464	24.2	570	1	AU040310
11	428.2	22.3	844	7	CK653343
12	397.4	20.7	735	4	B1756228
13	391.2	20.4	666	4	B1480798
14	389.2	20.3	639	6	CA389968
15	381.2	19.9	491	7	CU024947
16	368.4	19.2	582	5	BP379588
17	367	19.2	582	5	BP379588
18	342.2	17.9	508	2	BE236901
19	336	17.5	526	4	BM685501
20	334	17.4	335	5	BY351624
21	333.2	17.4	522	4	BM707649
22	327.6	17.1	819	7	CK408463
23	319.2	16.7	934	7	CF241451
24	310.6	16.2	497	6	CA391418

25	309	16.1	863	5	BU205267	BU205267
26	306	16.0	793	4	BG961794	BG961794
27	295.6	15.4	1067	4	BM562042	BM562042
28	291	15.2	1067	4	AV040568	AV040568
29	288.8	15.1	566	4	BM718338	BM718338
30	281.8	14.7	310	1	AV204954	AV204954
31	279.8	14.6	298	5	BK634624	BK634624
32	277.2	14.5	721	7	CN427090	CN427090
33	272.6	14.2	287	1	AV043872	AV043872
34	272	14.2	455	4	BG013943	BG013943
35	269.2	14.1	282	1	AV043522	AV043522
36	261.6	13.7	784	7	CF783151	CF783151
37	261	13.6	732	4	BG965806	BG965806
38	259.8	13.6	603	1	AA205892	AA205892
39	259.2	13.5	654	2	BE679896	BE679896
40	256.4	13.4	780	5	BX920526	BX920526
41	254	13.3	271	1	AA497726	AA497726
42	252.8	13.2	883	5	BU523270	BU523270
43	241.6	12.6	941	5	BU514796	BU514796
44	240.6	12.6	514	4	B1343182	B1343182
45	239.8	12.5	569	5	BP220982	BP220982

ALIGNMENTS

RESULT 1	AK006549	1699 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK006549				
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:1700030H21 product:vtc11iform macular dystrophy 2 homolog (human), full insert sequence.				
ACCESSION	AK006549	GI:12839710			
VERSION	AK006549.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriongnathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.				
TITLE	Normalizaton and subraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishii, K., Kitanai, T., Tashtiro, H., Itoh, M., Sumi, N., Iehi, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Watanabe, H., Sekiguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.				
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer				
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)				
MEDLINE	20530913				
PUBMED	11076861				
REFERENCE	4				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium				
TITLE	Functional annotation of a full-length mouse cDNA collection				
JOURNAL	Nature 409, 685-690 (2001)				
MEDLINE	5				
PUBMED	11076861				
REFERENCE	The FANTOM Consortium and the RIKEN Genome Exploration Research				

Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1699)
Adachi, Y., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
JOURNAL
REFERENCE
AUTHORS

Akawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiromoto, K., Hisaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Iwama, M., Kikukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Nunehara, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saio, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shimagawa, A., Shiraki, F., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT Please visit our web site (<http://genome.gsc.riken.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGACAGCATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGAGCGCCGCAATTATTCCTCAGTATTAATTAATATCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES	Location/Qualifiers
source	1..1699
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="PANTOM.DB:1700030H21"
	/db_xref="taxon:10090"
	/clone="1700030H21"
	/sex="male"
	/tissue_type="testis"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	1..1699
misc_feature	/note="Vitellogenin macular dystrophy 2 homolog (human) (MGD) [GI:1346332]"

ORIGIN			
Query Match	87.9%	Score 1684.8,	DB 3, length 1699,
Best Local Similarity	99.8%	Pred. No. 0;	
Matches 1697; Conservative	0;	Mismatches	2; Indels 1; Gaps 1

QY	179	GATCAGCAGCTGTTGTTGAGAGCGTGCTCTGTAAC	CGCAGCAAGCTTAACATTAAGCTC	AAC	238
Db	1	GATCAGCAGCTGTTGTTGAGAGCTGCTCTGTAAC	CGCAGCAAGCTTAACATTAAGCTC	AAC	60
QY	239	CCTATATCCTTCGTTCTGGGTTTCTATGTTA	CATTGGTGTGAGCCGCTGCTGAGCCAG		298
Db	61	CCTATATCCTTCGTTCTGGGTTTCTATGTTA	CATTGGTGTGAGCCGCTGCTGAGCCAG		120
QY	299	TACGAGAACTTCGCGTGCCCGACCGCTCAT	TGATCAGAGTGTCTAGCTTCGTGAGGGC		358
Db	121	TACGAGAACTTCGCGTGCCCGACCGCTCAT	TGATCAGAGTGTCTAGCTTCGTGAGGGC		180
QY	359	AAGAGTGAAGAAAGCCGTTTGCTCGCGCG	CAGCAGCTCTCATCCGCTAACCCCACTCCG	GGCCAA	418

Db	181	AAGATGAGAGGCGCGTTGCTGCGCGCAGCGCTATCCGTACGCCAATCCGTGGCCAA	240
QY	419	GTGCTCATCTGCGAGCATCAGACCTCGTCTAACAGCGCTTCCACTCTTGACAC	478
Db	241	GTGCTCATCTGCGAGCATCAGACCTCGTCTAACAGCGCTTCCACTCTTGACAC	300
QY	479	CTGCTGCTAGCAGGTTTATATGACCCATGGGGAACATTAACAGTTGCAAGTTGGCCCTA	538
Db	301	CTGGGTCTAGCAGGTTTATATGACCCATGGGGAACATTAACAGTTGCAAGTTGGCCCTA	360
QY	539	CCACAACAACATTTGCGGTGCGCCCTGGGCTGTGGTCTTGCCAACTTGTCATGAAGCCTAT	598
Db	361	CCACA-AACACATTTGCGGTGCGCCCTGGGCTGTGGTCTTGCCAACTTGTCATGAAGCCTAT	419
QY	599	CTTGAAGTTCGAATCCGGACAACCGTCTGTCTCAGACCTGATGAATGAGGTGTACT	658
Db	420	CTTGAAGTTCGAATCCGGACAACCGTCTGTCTCAGACCTGATGAATGAGGTGTACT	479
QY	659	TTGGCTACCTAGTGTGGAACGCTGATACCTAACGATTAAGATCCCATTTGGTATAC	718
Db	480	TTGGCTACCTAGTGTGGAACGCTGATACCTAACGATTAAGATCCCATTTGGTATAC	539
QY	719	ACACAGTGTGTGACAGTGTGATATACAGCTTTTCTTGTGATGCTTATCGGAGGAG	778
Db	540	ACACAGTGTGTGACAGTGTGATATACAGCTTTTCTTGTGATGCTTATCGGAGGAG	599
QY	779	TTTCTGAACCCAAACAAAGACTACCCAGGCCATGATGATCTGTGTGTGCTGTCTTC	838
Db	600	TTTCTGAACCCAAACAAAGACTACCCAGGCCATGATGATCTGTGTGTGCTGTCTTC	659
QY	839	ACAATCTCTGAATTTTATTTCTAACATGCGGTGCTGAAGAGTGGCAGACAGCTCATCAC	898
Db	660	ACAATCTCTGAATTTTATTTCTAACATGCGGTGCTGAAGAGTGGCAGACAGCTCATCAC	719
QY	899	CCCTTCCGGGAGAGCATGATGATTTTGAACATACTGATCATTTGACAGAACTTGAC	958
Db	720	CCCTTCCGGGAGAGCATGATGATTTTGAACATACTGATCATTTGACAGAACTTGAC	779
QY	959	GTGTCCCTGTGTCCGTGATGGGATGTGACACAGAACTTGCTCCCATGGAACGTACATG	1018
Db	780	GTGTCCCTGTGTCCGTGATGGGATGTGACACAGAACTTGCTCCCATGGAACGTACATG	839
QY	1019	TACTGGAACGAGGACGCGCTCAGCGCGCTTACACAGCTGTCTTCCAGGCTTTCGCGG	1078
Db	840	TACTGGAACGAGGACGCGCTCAGCGCGCTTACACAGCTGTCTTCCAGGCTTTCGCGG	899
QY	1079	CATTCCTCATGGGCTCCACCTTCAACATCATGCGCTTAAGAAAGAAATTAGAGCTTTGG	1138
Db	900	CATTCCTCATGGGCTCCACCTTCAACATCATGCGCTTAAGAAAGAAATTAGAGCTTTGG	959
QY	1139	TCAAAAGAGAGGCTGACACCGGATTAAGAAAGAGTGTGTATAGCAGCACCATAGGCTGC	1198
Db	960	TCAAAAGAGAGGCTGACACCGGATTAAGAAAGAGTGTGTATAGCAGCACCATAGGCTGC	1019
QY	1199	TTCTTAGACTGCAACCCAAACTACCATTTTCCCTTGAAAGCTTAAAGCCAAACTA	1258
Db	1020	TTCTTAGACTGCAACCCAAACTACCATTTTCCCTTGAAAGCTTAAAGCCAAACTA	1079
QY	1259	TTGTGTTCTAAGAACCCCGCTCTCGAAGGCGCAGTGTAAAGATGTGCAACCAAAAACAG	1318
Db	1080	TTGTGTTCTAAGAACCCCGCTCTCGAAGGCGCAGTGTAAAGATGTGCAACCAAAAACAG	1139
QY	1319	AAAGATGTTCTGAATTTAAAGGTTCTGACTTCTTGAATATGTGTTCAAGGTTTAAAGG	1378
Db	1140	AAAGATGTTCTGAATTTAAAGGTTCTGACTTCTTGAATATGTGTTCAAGGTTTAAAGG	1199
QY	1379	AGAGGCTCCATTGTGGGCCAAGGCAACCCAGCAACCCATCTATGTGCAAGTCAAGCACC	1438
Db	1200	AGAGGCTCCATTGTGGGCCAAGGCAACCCAGCAACCCATCTATGTGCAAGTCAAGCACC	1259
QY	1439	TCCAGTTCAAGACAGGTGATGTGGCTTTTCCACAGATTTACAAAGAAATCTGTCAATGAA	1498

Db 1260 TCCAGTTGACACAGGTGATGGCCCTTCACAGATTACCAAGAAATCTGTCACTGAAA 1319
Qy 1499 AAGAAACTGTGAGATTAACTTGAACATTCACAGAGAGCCACAGAACATCTTCAACAG 1558
Db 1320 AAGAAACTGTGAGATTAACTTGAACATTCACAGAGAGCCACAGAACATCTTCAACAG 1379
Qy 1559 CGCCGTTTGAACAGATGTCAACCAATATACAGGCTTAATGAAAGACATGACAGATCC 1618
Db 1380 CGCCGTTTGAACAGATGTCAACCAATATACAGGCTTAATGAAAGACATGACAGATCC 1439
Qy 1619 TATCCCTACAGGATGAAAGCTGGACCAACCTGTTCTATAGATGATGCTTCAACAGCC 1678
Db 1440 TATCCCTACAGGATGAAAGCTGGACCAACCTGTTCTATAGATGATGCTTCAACAGCC 1499
Qy 1679 TGGCCCTGACTTGAAGAGATGCCAGAGGCACTGACCCAGTCAAAAG3CACACAGAGAG 1728
Db 1500 TGGCCCTGACTTGAAGAGATGCCAGAGGCACTGACCCAGTCAAAAG3CACACAGAGAG 1559
Qy 1739 CGACACCCAGAGAGTGTGTTCACAGACAGTCTAGCATTAATCTAGAAACCAAGATACTT 1798
Db 1560 CGACACCCAGAGAGTGTGTTCACAGACAGTCTAGCATTAATCTAGAAACCAAGATACTT 1619
Qy 1799 AATAGTCTGCTGTAACCAACCTGTATTTTACGATCTTTCCAAACTAGAGATTAAATA 1858
Db 1620 AATAGTCTGCTGTAACCAACCTGTATTTTACGATCTTTCCAAACTAGAGATTAAATA 1679
Qy 1859 AACGGAATATTTCTTTAGG 1878
Db 1680 AACGGAATATTTCTTTAGG 1699

RESULT 2
BO879880 963 bp mRNA linear EST 16-AUG-2002
LOCUS BO879880
DEFINITION AGENCOURT 8241531 Lupsk1_dorsal_root_ganglion Homo sapiens CDNA
BO879880
BO879880.1 GI:22271888
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 963)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@dbp-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LHAM1563 row: 1 column: 08
High quality sequence scop: 623.
Location/Qualifiers
1. 963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupsk1_dorsal_root_ganglion"
/note="Vector: pCMV-Sport6 (Life Technologies); Site_1:
NotI; Site_2: SalI; CDNA made by oligo-ct priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGCTCCG-3' and

FEATURES
Source

1. 963
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6180559"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupsk1_dorsal_root_ganglion"
/note="Vector: pCMV-Sport6 (Life Technologies); Site_1:
NotI; Site_2: SalI; CDNA made by oligo-ct priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGCTCCG-3' and

5'-GACTAGTTTATGATCGCAGCGCCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

ORIGIN

Query Match 29.9%; Score 573.6; DB 5; Length 963;
Best Local Similarity 83.4%; Pred. No. 5.2e-139;
Matches 651; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy 351 TGAAGGCGAAGATGAGAGAGCCGTTGCTGCGCGCAGCCTCATCGCTACGCCATCC 410
Db 26 TCGAAGGCGAAGATGAGAGAGCCGCTGCTGCGCGCAGCCTCATCGCTACGCCAACC 85
Qy 411 TGGGCGAAGTCTCATATCTGGCGCATCAGCACTCGGTCTACAGAGGCTTTCCCATC 470
Db 86 TGGGCGAAGTCTCATATCTGGCGCATCAGCACTCGGTCTACAGAGGCTTTCCCATC 145
Qy 471 TTCAACACCTGCTGAGAGGTTTATGACCCATGAGGGAATTAAGCAGTTGACAGAGT 530
Db 146 CCGACACCTGCTGAGAGGTTTATGACCTCCGCGAAGACAGAGAGTTGAGAGAAC 205
Qy 531 TGGGCGTACCAACACACATTTGCGGTCCTGGGCTGGTTGCTTCCAACTTGTCAATGA 590
Db 206 TGAAGCTTACCAACACACATTTGCGGTCCTGGGCTGGTTGCTTCCAACTTGTCAATGA 265
Qy 591 AGGCTTATCTTGGAGGTCGAATCCGGGACACCGTCTGCTCAAGACCTGATGAATGAGG 650
Db 266 AGGCTTATCTTGGAGGTCGAATCCGGGACACCGTCTGCTCAAGACCTGCTGAACGAGA 325
Qy 651 TGTGATCTTGGAGTCTGAGTGTGAGACAGCTGATAGCTTACAGTATGATATCCCAT 710
Db 326 TGAACACCTTGGAGTCTGAGTGTGAGACAGCTGATAGCTTACAGTATGATATCCCAT 385
Qy 711 TGTGATACACACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 770
Db 386 TGTGATACACACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 445
Qy 771 GGAGGAGTTTGTGAAACCCAAAGAGATACACAGGCTGATGATGATGATGATGATGATG 830
Db 446 GGAGGAGTTTGTGAAACCCAAAGAGATACACAGGCTGATGATGATGATGATGATGATG 505
Qy 831 CTGCTTACACATCTGCAATTTCTTATCTGATGAGGCTGATGAGTGTGAGTGTGAGTGTG 890
Db 506 CCGTCTTACACATCTGCAATTTCTTATCTGATGAGGCTGATGAGTGTGAGTGTGAGTGTG 565
Qy 891 TCAATCAACCCCTTGGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 950
Db 566 TCAATCAACCCCTTGGGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 625
Qy 951 ACTGCAAGTGTCTCTGTTGCTGCGTGAATGAGATGACACAGAACTTGCTTCCATGAGAC 1010
Db 626 ATTGCAAGTGTCTCTGTTGCTGCGTGAATGAGATGACACAGAACTTGCTTCCATGAGAC 685
Qy 1011 GTGACATGATGAG 1070
Db 686 CGGACATGATGAG 745
Qy 1071 CTCGCCGCAATTCCTTATGAGGCTTCAACCTTCAACATCAGCCTTAAGAGAGAGAGAGAG 1130
Db 746 TCCGTCAGAGCTCTTTATGAGGCTTCAACCTTCAACATCAGCCTTAAGAGAGAGAGAG 805
Qy 1131 A 1131
Db 806 A 806

RESULT 3
BY714331 713 bp mRNA linear EST 17-DEC-2002
LOCUS BY714331
DEFINITION BY714331 RIKEN full-length enriched, adult male testis Mus musculus

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS

cDNA clone 4921504J18 5' mRNA sequence.
 BY1414331
 BY1414331.1 GI:27127441
 EST.
 Mus musculus
 Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (bases 1 to 713)
 Okazaki, Y., Furuno, M., Kasukawa, T., Aachai, V., Bono, H., Kondo, S., Nikiado, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldorelli, R., Hill, D.P., Bule, C., Hune, D.A., Quackenbush, J., Schriml, L.H., Kapapin, A., Matsuda, H., Batalov, S., Beisel, K.M., Blake, J.A., Brad, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Drasgini, T.A., Fletcher, C.F., Forrest, A., Fraser, K.S., Gaasterland, T., Gaitoldi, C.M., Gissi, C., Godkaj, A., Gough, J., Grimmond, S., Gustlin, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kani, A., Kawai, H., Kawasawa, Y., Kedierski, R.M., King, B.L., Kondagala, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semp, C.A., Secou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yangisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kitahara, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shitaki, T., Waki, K., Kawai, J., Aizawa, K., Akakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Miyazaki, A., Sakai, R., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, B.S., Rongers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 22354683
 Nature 420, 563-573 (2002)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Yoshitake Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Saito-cho
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.jp, URL: http://genome-gsc.riken.jp/
 Adachi, J., Aizawa, K., Akimura, T., Akakawa, T., Caminci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Horii, P., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Kawai, J., Kojima, Y., Komura, S., Kono, H., Koyama, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Osato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shitaki, T., Tagami, T., Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopaedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES	location/Qualifiers
source	1..713

```

/tissue_type="testis"
/dev_stage="adult"
/lab_host="SOLR"
/clone_1ib="RIKEN full-length enriched, adult male testis
/notes="Site 1: XhoI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGGAGAGAGAGATCCAGAGCGCTCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGGAGAGAGCGCGCGCATTTATCTTCAGGTTATTAATTATCCCCCCCCC
3']. cDNA was cloned into the XhoI and BamHI sites."

```

Query Match	27.1%	Score 519;	DB 6;	Length 713;
Best Local Similarity	97.8%;	Pred. No. 1e-124;		
Matches 525; Conservative	0;	Mismatches 12;	Indels 0;	Gaps 0

QY	1	GTGGCAAGCCATAGACATATACACCTCAACAACAAAGTATACCAATCCCGGCTCCGGTTGGTT	60
Dp	177	GTGGCAAGCCATATACATATACCTTACACAAACAAATACCAATCCCGGCTCCGGTTGGTT	236
QY	61	CTGGTCCCTCTCTCTGTGTGGCGAAGGACATCTACAGCTGGCTGTATGAGAAATTCTT	120
Dp	237	CTGGTCCCTCTCTCTGTGTGGCGAAGGACAGATCTACAGCTGGCTGTATGAGAAATTCTT	296
QY	121	TGTCTTCATATTCCTCTACTATTTTCATCCGTGGAGCTACAGCAATGGTTCTCTCGAGTA	180
Dp	297	TGTCTTCATATTCCTCTACTATTTTCATCCGTGGAGCTACAGCAATGGTTCTCTCGAGTA	356
QY	181	TCAGCAAGCTGTGTGTTGAGAAAGCTGGCTCTGTACTGTGAGACAGTACATTCAGCTCATCC	240
Dp	357	TCAGCAAGCTGTGTGTTGAGAAAGCTGGCTCTGTACTGTGAGACAGTACATTCAGCTCATCC	416
QY	241	TATATCTCTCGTTCCGGGTTCTAATGTAACTATGGTGTGAGCCGCTGGTGGAGCCAGTA	300
Dp	417	TATATCTCTCGTTCCGGGTTCTAATGTAACTATGGTGTGAGCCGCTGGTGGAGCCAGTA	476
QY	301	CGAGAACTTGGCGTGGCCGACCGGCTCATAGTACAGGTGTCTAGGTTCTGTGAGGGCA	360
Dp	477	CGAGAACTTGGCGTGGCCGACCGGCTCATAGTACAGGTGTCTAGGTTGTGAGGGCA	536
QY	361	GGATGAGAAAGGCGTTTGTGTGGGGCCAGGCTCATTCGCTTAGCCATCTCTGGGCCAAGT	420
Dp	537	GGATGAGAAAGGCGTTTGTGTGGGGCCAGGCTCATTCGCTTAGCCATCTCTGGGCCAAGT	596
QY	421	GCTCATCTGTGCGAGCATCAGCACTTGGTCTCAAGCGCTTTCACACTTTCACACACT	480
Dp	597	GCTCATCTGTGCGAGCATCAGCACTTGGTCTCAAGCGCTTTCACACTTTCACACACT	656
QY	481	GGGTCTAGCAGGTTTATGACCCATGGGGAACTAAGCAAGTGTGGAGCT	537
Dp	657	GGGTCTAGCAGGTTTATGACCCATGGGGAACTAAGCAAGTGTGGAGCT	713

RESULT 4
CD518675

LOCUS CD518675 881 bp mRNA linear EST 06-JUN-2003
 DEFINITION AGNCOURT 14375996 NIH_MGC_181 Homo sapiens cDNA clone
 IMAGE:3037964 5', mRNA sequence.
 ACCESSION CD518675
 VERSION CD518675.1 GI:31450393
 KEYWORDS EST.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 881)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE NIH-MGC Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
 http://image.llnl.gov
 Plate: NDAM482 row: b column: 21
 High quality sequence stop: 652.
 Location/Qualifiers
 1..881
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3037964"
 /tissue_type="White Matter"
 /dev_stage="Unknown"
 /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
 /clone_1lb="NIH_MGC_181"
 /note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV
 (destroyed); Library is oligo-AT primed and directionally
 cloned (EcoRV site is destroyed upon cloning). Average
 insert size 1.42 kb. Library was constructed by
 (Invitrogen). Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 26.2%; Score 502.6; DB 6; Length 881;
 Best Local Similarity 80.1%; Pred. No. 2.2e-120;
 Matches 608; Conservative 0; Mismatches 139; Indels 12; Gaps 1;

QY 476 CACCTGGTGTGACGAGTTTATGACCATGCGGAGCAATAGCACTTCCAGAACTTGGGC 535
 Db 1 CACCTGGTGTGACGAGTTTATGACCATGCGGAGCAATAGCACTTCCAGAACTGAGC 60
 QY 536 CTACACACAAACATTTCTGGGTGCTGCTGGGTGTGTTGCCAATTTTCATATGAAGCC 595
 Db 61 CTACACACAAACATTTCTGGGTGCTGCTGGGTGTGTTGCCAATTTTCATATGAAGCC 120
 QY 596 TATCTGGAGGTGCAATCCGGACACCGTCTGCTCCAGAGCTGATTAATGAGGTGTGT 655
 Db 121 TGGCTTGGAGGTGCAATCCGGACACCGTCTGCTCCAGAGCTGATTAATGAGGTGTGT 180
 QY 656 ACTTTGGCTACTGAGTGTGACAGCTGTATGCTTATGCTGATTAAGTATCCATTTGGTG 715
 Db 181 ACTTTGGCTACTGAGTGTGACAGCTGTATGCTTATGCTGATTAAGTATCCATTTGGTG 240
 QY 716 TACACACAGTGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 775
 Db 241 TATACACAGTGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 300
 QY 776 CAGTTTTCGAACCCCAAGAGACTACCGAGGCCATGAGATGATCTGTTGTGCTGTC 835
 Db 301 CAGTTTTCGAACCCCAAGAGGCCATACCTGCGCATGAGCTGAGACTGTTGTGCTGTC 360

QY 836 TTCAACATCTGCAATTTCTTATTTACATAGGAGCTGAGTGAAGTGGCAGAAACAGCTCATC 895
 Db 361 TTCAAGTTCCTGCAAGTTTCTTTCTATGTTGGCTGGCTGAAAGTGGCAGAGCACTCATC 420
 QY 896 AACCCCTTGGGAGAGACGATGATGATTTTGAAGTAAGTGGATATTGACAGAAACCTG 955
 Db 421 AACCCCTTGGAGAGATGATGATGATTTTGAAGTAAGTGGATATTGACAGAAATTTG 480
 QY 956 CAGGTGCTCCGTTGTGCTGGTGGATGAGTGAACCGAAGCTTGGCCCATGGAGAGTGC 1015
 Db 481 CAGGTGCTCCGTTGTGCTGGTGGATGAGTGAACCGAAGCTTGGCCCATGGAGAGTGC 540
 QY 1016 ATGTACTGGAACGAGGACCGCTCAGCGCCCTTACAGAGTGTCTTTCAGAGTCTGC 1075
 Db 541 ATGTACTGGAATTAAGCCCGACAGAGCCCTTACAGAGTGTCTTTCAGAGTCTGC 600
 QY 1076 CGGACTTCTTCAATGAGGCTCCACTTCAACATCAGCTTAAAGAAAGACTTAGACTT 1135
 Db 601 CGAGCTCTCTTATGAGGCTCCACCTTCAACATCAGCTTAAAGAAAGAGTGAAGTTC 660
 QY 1136 TGTCTAAAGAGAGGCTGACACGATTAAGAAAGAGTGGCTATAGCAGACCATAGG 1195
 Db 661 CAGCCCAATCAGAGG-----ACAGAGAGATGCTCAGCTGGCATCATTTGAC 708
 QY 1196 TGTCTTATGAGACTGCAACCCAAACATCACTATCTCC 1234
 Db 709 CGTTTCTTATGAGCTGCACTCCATGATCACTATCTCC 747

RESULT 5
 CK470840 792 bp mRNA linear EST 13-JAN-2004
 LOCUS AGNCOURT 17639671 NIH_MGC_237 Rattus norvegicus cDNA clone
 DEFINITION IMAGE:7115295 5', mRNA sequence.
 ACCESSION CK470840
 VERSION CK470840.1 GI:40809322
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 792)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Howard Jacobs
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
 http://image.llnl.gov
 Plate: LLAM1490 row: d column: 13
 High quality sequence stop: 648.
 Location/Qualifiers
 1..792
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7115295"
 /tissue_type="testis, pooled"
 /lab_host="DH10B-TonA"
 /clone_1lb="NIH_MGC_237"
 /note="Organ: testis; Vector: pExpress-1; Site_1: EcoRV;
 Site_2: NotI; RNA obtained from testis tissue of 8 wk old
 animal. Tissues were snap-frozen and kept at -80C before
 RNA extraction and purification (TRI-reagent method). cDNA

was primed using oligo-dT primer:
5'-pGACTAGTTCAGATCGGAGCGGCCGCC(T)25-3' and cloned into
the EcoRV/NotI sites of pEXpress-1. Size-selection >1.4kb
resulted in an average insert size of 2.4 kb. This primary
library is not normalized (normalized primary library is
NIH MGC 238) and was constructed by Express Genomics
(Frederick, MD)"

ORIGIN

Query Match 25.8%; Score 493.6; DB 7; Length 792;
Best Local Similarity 93.5%; Pred. No. 5e-118;
Matches 536; Conservative 0; Mismatches 35; Indels 2; Gaps 2;

```
QY 1 GTTCCAGGCACTATCTATCACTACCAACAAAGTAGCCATGCCGCTCGTTCCTT 60
DB 142 GTGCCAAGCCATGACTATCACTACCAACAAAGTAGCCATGCCGCTCGTTCCTT 201
QY 61 CTGTGCTCTCTCTCTGCTGGGAGGAGGAGCATCTCAAGCTCTGTATGAGAAATTCCT 120
DB 202 CTGTGCTCTCTCTCTGCTGGGAGGAGGAGCATCTCAAGCTCTGTATGAGAAATTCCT 261
QY 121 TGTCTTCAATATCTCTATCTATCTATCCATCCGTGAGCTCTCAAGATGTTCTCTCGAGTGA 180
DB 262 TGTCTTCAATATCTCTATCTATCTATCCGTGAGCTCTCAAGATGTTCTCTCGAGTGA 321
QY 181 TCAGCAGCTGTTGTTGAAGAGCTGCTCTGTACTGCGACAGCTTACATTCAGCTTATCCC 240
DB 322 TCAGCAGCAGCAGTTGAGAAAGTTGGCTCTGTACTGCAACAGCTTACATTCAGCTTATCCC 381
QY 241 TATATCTCTGATCTGAGTTTCTATGTTCAATTTGGTGGTGAAGCCGCTGTGAGGAGCA 300
DB 382 TATATCTCTGATCTGAGTTTCTATGTTCAATTTGGTGGTGAAGCCGCTGTGAGGAGCA 441
QY 301 CGAAGACTTCCGCTGAGCCGACCGCTCAATGATCCAGATGTTAGCTTCTGAGAGGCA 360
DB 442 CGAAGACTTCCGCTGAGCCGACCGCTCAATGATGAGGATGTTAGCTTCTGAGAGGCA 501
QY 361 GGATGAGGAGAGGCGCTTGTGCTGGGAGGAGCGCTCATCCGCTAGCCATCTGAGGCAAGT 420
DB 502 GGAGCAGGAGAGGCGCTTGTGCTGGGAGGAGCGCTCATCCGCTAGCCATCTGAGGCAAGT 561
QY 421 GCTCATCTCTGCGACATCAGCAGCTCGGTCTACAAAGGCGCTTCCACATCTTCAACCACT 480
DB 562 GCTCATCTCTGCGACATCAGCAGCTCGGTCTACAAAGGCGCTTCTCACTTCTCAACCACT 621
QY 481 GGTCTAGCAGAGTTTATGACCATGCGGAGAACATTAAGCAGTTGCAAGTTGGGCTTACC 540
DB 622 GATGCAACGAGCTTATGACTCTGNGAGAAATA-AAGCAGTTGCAAGGTTGGGCTTACC 680
QY 541 ACACAACACATTCGGGAGCGCTGAGTGGT 573
DB 681 ACACAACAC-TTCTGGGTACCTGGGTTGGTT 712
```

RESULT 6
LOCUS CO396179 756 bp mRNA linear EST 01-JUL-2004
DEFINITION AGENCOURT 2782871 NIH_MGC_212 Homo sapiens cDNA clone
IMAGE:30923539 5', mRNA sequence.
ACCESSION CO396179
VERSION CO396179.1 GI:49578095
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH

FEATURES

source

BLdg, 31 Rm10A07 Bethesda, MD 20892
Email: cga@bldg.nih.gov
Tissue Procurement: Mary Hendrix
cDNA Library Preparation: M. Bento Soares, University of Iowa
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lim.gov
Plate: NDM1164 row: 1 column: 20
High quality sequence stop: 571.
Location/Qualifiers
1..756

ORIGIN

Query Match 25.5%; Score 488.2; DB 7; Length 756;
Best Local Similarity 84.1%; Pred. No. 1.3e-116;
Matches 550; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

```
QY 323 CGCTCATGATCCAGAGTGTCTAGCTTCTGAGAGGAGGAGATGAGAGGCGCTTGTCTG 382
DB 8 CGCTCATGAGCTGTGGTGTGGGCTTCTGAGAGGAGGAGGAGGAGGCGCTGTCTG 67
QY 383 CGGCGCAGCTTATTCGCTACAGCCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 442
DB 68 CGGCGCAGCTTATTCGCTACAGCCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 127
QY 443 ACCTCGCTTACAGGAGCTTCCACTTCTTCAACCACTGCTGTGACAGTCTTATGAGC 502
DB 128 ACCTCGCTTACAGGAGCTTCCAGCGCCAGCAGCTGTGTGAGCAGGCTTATGAGC 187
QY 503 CATGGGAGAACATTAAGCAGTTGAGAAAGTTGGGCTTACACACACACATTCGTGGTCCC 562
DB 188 CCGGAGAGAACATTAAGCAGTTGAGAAAGTTGGGCTTACACACACATTCGTGGTCCC 247
QY 563 TGGGTGGTGGTGGGCAACTTGTCAATGAGAGGCTTATCTTGGAGGTGGAATCCGGAGCAC 622
DB 248 TGGGTGGTGGTGGGCAACTTGTCAATGAGAGGCTTATCTTGGAGGTGGAATCCGGAGCAC 307
QY 623 GTCTGCTCTCAGAGCTGATGATGAGGAGTGTATCTTGTGCTGCTCAAGTGGAGAGCTG 682
DB 308 ATCTGCTCTCAGAGCTGCTGATGATGAGGAGTGTATCTTGTGCTGCTCAAGTGGAGAGCTG 367
QY 683 TATGCTTACGACTGATTAAGTATCCATGTTGTGACACACAGGTGTGACAGTGGACAGTA 742
DB 368 TATGCTTACGACTGATTAAGTATCCATGTTGTGACACACAGGTGTGACAGTGGACAGTA 427
QY 743 TACAGCTTCTCTGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 802
DB 428 TACAGCTTCTCTGCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 487
QY 803 CCAGGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 862
DB 488 CCGGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 547
```


Oy	863	ATGGCGTCGCTGAAGTGTGGCAAGAAGCATTAACAACCCCTTGCGGAGAGAGATGATAT	922
Db	548	GTTGGCTTGGCTGAAAGTAGCAGAGACGCTCATCAACCCTTTGGAGAGAAATGATGATAT	607
Oy	923	TTTGAGACTAATCGATCATTTGACAGAAACCTGCAGAGTGTCCCTATTGCCGTG	976
Db	608	TTTGAGACCAATTGGATTGTTCGAAAGAAATTTCGAAAGTGTCCCTGTGCTGTG	661
RESULT 7			
LOCUS	BY714344	664 bp	mRNA linear EST 17-DEC-2002
DEFINITION	BY714344 RIKEN full-length enriched, adult male testis Mus musculus		
ACCESSION	CDNA clone 4921506F15 5', mRNA sequence.		
VERSION	BY714344		
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 664)		
REFERENCE	Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Oseo,N., Saito,K., Tomaru,Y., Suzuki,H., Yamana,T., Kiyosawa,H., Yagi,K., Tomaru,Y., Haegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.F., Butt,C., Hume,D.A., Quackenbush,J., Schirral,L.M., Kanapin,A., Matsuda,H., Beralov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusci,V., Chochla,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frizer,K.S., Gaasterland,T., Gariboldi,M., Glasei,C., Godzi,A., Gough,J., Grimmond,S., Guilicinski,S., Hirokawa,N., Jackson,I.J., Jarvis,R.D., Kanai,A., Kawai,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Kongaya,A., Kurouchkin,I.V., Lee,Y., Lehnard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,M.J., Petrea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shmada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M., Verardo,R., Wagner,L., Wahlesedt,C., Wang,Y., Waanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yangisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Saeki,D., Shibata,K., Shinagawa,J., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,A. and Hayashizaki,Y.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
PUBMED	22354683		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-research.riken.jp, URL:http://genome.gsc.riken.jp/ Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,T., Kojima,Y., Konodo,S., Konno,H., Koys,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Saeki,D., Sato,K., Shibata,K., Shiraki,T., Takami,M., Takeda,Y., Waki,K., Watanahi,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to		

FEATURES	SOURCE
prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)	
RIKEN Integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)	
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)	
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.	
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.	
Location/Qualifiers	
1..664	
/organism="Mus musculus"	
/mol_type="mRNA"	
/strain="C57BL/6J"	
/db_xref="taxon:10090"	
/clone="4921506F15"	
/sex="male"	
/tissue_type="testis"	
/dev_stage="adult"	
/lab_host="SOLR"	
/clone_id="RIKEN full-length enriched, adult male testis"	
/note="Site 1: XhoI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5].	
GAGGAGAGAGAGATCCAGAGCTTTTTTTTTTTTTNN 3'). cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5].	
GAGGAGAGAGAGCGGCGGCATTAATTCAGTTAATTAATTCGCCGCC 3'). cDNA was cloned into the XhoI and BamHI sites. "	
ORIGIN	
Query Match 25.4%; Score 487; DB 6; Length 664;	
Best Local Similarity 99.8%; Pred. No. 2.6e-116;	
Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
1 GTGCCAAGCCATGACATCACTACACAAACAAAGTACCAATGCCGCTCGGTTGCTT 60	
177 GTGCCAAGCCATGACATCACTACACAAACAAAGTACCAATGCCGCTCGGTTGCTT 236	
61 CTGCTCCCTCCTCGTGGCTGGGAGGACAGATCTAAGCTGCTGATGAGAAATCCT 120	
237 CTGATCCCTCCTCGTGGCTGGGAGGACAGATCTAAGCTGCTGATGAGAAATCCT 296	
121 TGTCTTCAATATTCCTACTATTTCATCCGTCGACTTACAGAAATGTTCTCTGATGA 180	
297 TGTCTTCAATATTCCTACTATTTCATCCGTCGACTTACAGAAATGTTCTCTGATGA 356	
181 TCAGCAGCTGTGTTTGAAGAGCTGAGCTGTGATCTGCGAGAGCTACATTGACGTCATCCC 240	
357 TCAGCAGCTGTGTTTGAAGAGCTGAGCTGTGATCTGCGAGAGCTACATTGACGTCATCCC 416	
241 TATATCCTTGTGTTGGGTTTCTATATTGAATGGTGTGACCGCTGGTGGAGCCAGTA 300	
417 TATATCCTTGTGTTGGGTTTCTATATTGAATGGTGTGACCGCTGGTGGAGCCAGTA 476	
301 CGAAGAACTTCCGTGGGCGCAGCCGCTCATGATCCAGGTGTTAGCTTCGTGAGGAGCA 360	
477 CGAAGAACTTCCGTGGGCGCAGCCGCTCATGATCCAGGTGTTAGCTTCGTGAGGAGCA 536	
361 GGATGAGAAAGCGGTTGCTGGGCGCAGCTTCATCCGTAAGCCATCTCTGGGCCAAGT 420	

Db 537 GGATGAGAAAGCGCGTTTCTGCGGGGAGACGCTATCCGCTACCGCATCTCGGCCAAGT 596
 Qy 421 GGTATCTCGCGGACATCAGACCTGGTCTTACAGAGCGTTTCCACTTTCAACCACT 480
 Db 557 GGTATCTCGCGGACATCAGACCTGGTCTTACAGAGCGTTTCCACTTTCAACCACT 656
 Qy 481 GGTCCTAG 488
 Db 657 GGTCCTAG 664

RESULT 8
 LOCUS CK469826 862 bp mRNA linear EST 13-JAN-2004
 DEFINITION AGNCOURT 17644559 NIH MGC 237 Rattus norvegicus cDNA clone
 IMAGE:7113326 5', mRNA sequence.

ACCESSION CK469826
 VERSION CK469826.1 GI:40808308
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 862)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@os-remail.nih.gov
 Tissue Procurement: Howard Jacobs
 CDNA Library Preparation: Express Genomics
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILMIL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILMIL at:
 http://image.llnl.gov
 Plate: LLM14985 row: b column: 12
 High quality sequence stop: 674.
 Location/Qualifiers

FEATURES
 source 1..862

/organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7113326"
 /tissue_type="testis, pooled"
 /lab_host="DH10B Tona"
 /clone_lib="NIH MGC 237"
 /note="Organ: testis; Vector: pExpress-1; Site: 1: EcoRV;
 Site 2: NotI; RNA obtained from testis tissue of 8 wk old
 animal. Tissues were snap-frozen and kept at -80C before
 RNA extraction and purification (TRI-reagent method). cDNA
 was primed using oligo-dT primer:
 5'-GGACTAGTTCTAGATCGGAGCGCGCC(T)25-3' and cloned into
 the EcoRV/NotI sites of pExpress-1. Size selection >1.4kb
 resulted in an average insert size of 2.4 kb. This primary
 library is not normalized (normalized primary library is
 NIH-MGC 238) and was constructed by Express Genomics
 (Frederick, MD)"

ORIGIN

Query Match 24.7%; Score 474.2; DB 7; Length 862;
 Best Local Similarity 92.0%; Pred. No. 6.3e-113;
 Matches 532; Conservative 0; Mismatches 41; Indels 5; Gaps 3;

Qy 1 GTGCGAAGCGATGATACCTACCTACACAAACAAAGTAGCAATGCCGCTGGTTCGT 60
 Db 158 GTGCGAAGCGATGATACCTACCTACACAAACAAAGTAGCAATGCCGCTGGTTCCT 217
 Qy 61 CTGCTCCCTCTCTGTGCTGGAGGAGGAGCATCTCAAGCTGCTGTATGAGAAATTCCT 120

Db 218 CTGAGGCTCTCTCGGCGCTGGCGAGGACATCTCAAGCTGTGACGAGAAATTCCT 277
 Qy 121 TGTCTTCAATATTCCTCTACATATTCATCGTGAAGCTACAGATGTTCTCTGAATGA 180
 Db 278 TGTCTTCAATATTCCTCTACATATTCATCGTGAAGCTACAGATGTTCTCTGAATGA 337
 Qy 181 TCAGAGCTGTGTTTGAAGCTGAGCTGCTACTGCTGAGACAGCTACATTCAGTCAATCC 240
 Db 338 TCAGAGAGACAGATTGAGAGATTGAGCTCTGACTGCAAGAGCTACATTCAGTCAATCC 397
 Qy 241 TATATCTTCTGTTCTGAGTTTCTATATGTTACATTTGTTGAGAGCGCTGTGAGCCAGTA 300
 Db 398 TATATCTTCTGTTCTGAGTTTCTATATGTTACATTTGTTGAGAGCGATGTTGAGAACAGTA 457
 Qy 301 CGAAGACTGCGCGAGGCGGACCGCTCAAGATCCAGTGTCTTACCTTGTGAGAGGCA 360
 Db 458 CGAAGACTGCGCGAGGCGGACCGCTCAAGATCCAGTGTCTTACCTTGTGAGAGGCA 517
 Qy 361 GGATGAGAAAGCGCGTTTGTGCGGCGACGCTCATCCGTCAGCCATCTGGGCGCAAGT 420
 Db 518 GGAAGAGCAAGCGCGGTTGCTGCGGCGACGCTCATCCGTCAGCCATCTGGGCGCAAGT 577
 Qy 421 GGTATCTCGCGGACATCAGACCTGGTCTTACAGAGCGTTTCCACTTTCAACCACT 480
 Db 578 GGTATCTCGCGGACATCAGACCTGGTCTTACAGAGCGTTTCCACTTTCAACCACT 637
 Qy 481 GGTCCTAGAGTTTATGACCCATGAGGAGATAGAGTGTGAGAGTGGGCTACC 540
 Db 638 GATGCAAGCAGGCTTTATGACTCTGNGAATTA-AGCAGTTGCAAGAGTGGGCTTACC 696
 Qy 541 ACACACACATTTTGGTGCGCTGCGCTGGTGTGTTGCCA 578
 Db 697 AC--ACACATTTGGGT-ACCTGGGTGTGNTGCNA 730

RESULT 9
 LOCUS AU040308/c 578 bp mRNA linear EST 04-DEC-1998
 DEFINITION AU040308 Mouse four-cell-embryo cDNA Mus musculus cDNA clone
 J0803G03 3', mRNA sequence.
 ACCESSION AU040308
 VERSION AU040308.1 GI:3954128
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 578)
 AUTHORS Ko, M.S.H., Kitchen, J.R., Wang, X., Threat, T.A., Sun, T.,
 Depalma, G.E., Liang, Y., Kargul, G.J., Shatara, R., Lim, M.K. and
 Doi, H.
 TITLE Systematic analyses of genes expressed in 4-cell mouse embryo (The
 ERAO/Doi Project at Wayne State University)
 JOURNAL Unpublished (1998)
 COMMENT Contact: Hirofumi Doi
 Doi Bioscience Project, ERAO
 Japan Science and Technology Corporation (JST)
 WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan
 Email: hdoibioa.jst.go.jp.
 Location/Qualifiers

FEATURES
 source 1..578

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="J0803G03"
 /dev_stage="four-cell-embryo"
 /clone_lib="Mouse four-cell-embryo cDNA"

ORIGIN

Query Match 24.3%; Score 466; DB 1; Length 578;
 Best Local Similarity 94.6%; Pred. No. 8.2e-111;

[illegible]

ORIGIN	Query Match	Best Local Similarity	24.2%: Score 464; DB 1; Length 570;	Matches 532; Conservative 0; Mismatches 23; Indels 6; Gaps 5;
			/clone="J0803G08"	
			/dev_stg="four-cell-embryo"	
			/clone_id="Mouse four-cell-embryo cDNA"	
QY	1321	AGATGTCGAAATTTAAAGGCTGCACTTCTTGAATGTGTTCCAGGTTTAAGAGAG	1380	
DB	556	AGAAAGATCTGAATTTAAGGCTCTGACCTTC-TGAATGTGTTCCAGGTTTAAGAGAG	500	
QY	1381	AGGCTCCCATTTGTGGCCCAACAGGACCCACAGGACCTTACGTAGAGAGTACAGCCCTC	1440	
DB	499	AGGCTCCCATTTGTGGCCCAACAGGACCCACAGGACCTTACGTAGAGAGTACAGCCCTC	441	
QY	1441	CAG-TTCAGACACAGGTGATGGGCTTCCACAGTTACCAAGAAATCTGTCACTGAAAA	1499	
DB	440	CAGTTTCAGACACAGGTGATGGGCTTCCACAGTTACCAAGAAATCTGTCACTG-AAA	382	
QY	1500	AGAAACTGTGGAATTTAACTTTGAACATTCAGAGAGCCCAACAGACATCTTCAACAGC	1559	
DB	381	AGAAACTGTGGAATTTAACTTTGAACATTCAGAGAGCCCAACAGACATCTTCAACAGC	322	
QY	1560	GCCGTTTGACACAGATGTCAACCAATTAACAGGCTCTAATGAGAGACATGACAGTCT	1619	
DB	321	GCCGTTTGACACAGATGTCAACCAATTAACAGGCTCTAATGAGAGACATGACAGTCT	262	
QY	1620	ATCCCTACAGGAGATGAAAGCTGGACCAACCTGTCTTATGAGTATGCTTCAACAGCT	1679	
DB	261	ATCCCTACAGGAGATGAAAGCTGGACCAACCTGTCTTATGAGTATGCTTCAACAGCT	202	
QY	1680	GAGCCTGACTTGCAGAGATGGCCAGGAGGACCTGACCCAGTCAAGAGACACAGAGAGC	1739	
DB	201	GAGCCTGACTTGCAGAGATGGCCAGGAGGACCTGACCCAGTCAAGAGACACAGAGAGC	142	
QY	1740	GACACCCAGAGAGTGTGTTCCACAGCAGTCTAGCATGTAACTCAGAACCAAGACTCTTA	1799	
DB	141	GACACCCAGAGAGTGTGTTCCACAGCAGTCTAGCATGTAACTCAGAACCAAGACTCTTA	82	
QY	1800	ATAGTCTGCTGTAAAAACCTGTATTTTACATCTTTCCAAACTAAGAGTTTAATA	1859	
DB	81	ATAGTCTGCTGTAAAAACCTGTATTTTACATCTTTCCAAACTAAGAGTTTAATA	22	
QY	1860	ACGTGAATATCTTTTAAAGTG 1880		
DB	21	ACGTGAATATCTTTTAAAGTG 1		
RESULT 11				
LOCUS	CK653343	844 bp	mRNA	linear EST 29-JAN-2004
DEFINITION	AGNCOCURT_1674107 NIH_MGC_237 Rattus norvegicus cDNA clone			
IMAGE:	7114591 5', mRNA sequence.			
ACCESSION	CK653343			
VERSION	CK653343.1 GI:41388710			
KEYWORDS	EST.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
	Rattus.			
REFERENCE	1 (bases 1 to 844)			
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/.			
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL	Unpublished (1999)			
COMMENT	Contact: Daniela S. Gerhard, Ph.D.			
	Office of Cancer Genomics			
	National Cancer Institute / NIH			
	Bldg. 31 Rm10A07 Bethesda, MD 20892			
	Email: cgabs-remail.nih.gov			
	Tissue Procurement: Howard Jacobs			

CDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM14988 row: 9 column: 05
High quality sequence stop: 669.
Location/Qualifiers

FEATURES
source

1..844
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="IMAGE:7114591"
/issue_type="testis, pooled"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_237"
/note="Organ: testis; Vector: pExpress-1; Site_1: EcoRV;
Site_2: NotI; RNA obtained from testis tissue of 8 wk old
animal. Tissues were snap-frozen and kept at -80C before
RNA extraction and purification (Tri-reagent method). CDNA
was primed using oligo-dT primer:
5'-pGACTGTTCTAGATGCGAGCGGCCGCC(T)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size selection >1.4kb
resulted in an average insert size of 2.4 kb. This primary
library is not normalized (normalized primary library is
NIH_MGC_238) and was constructed by Express Genomics
(Frederick, MD)"

ORIGIN

Query Match 22.3%; Score 428.2; DB 7; Length 844;
Best Local Similarity 95.1%; Pred. No. 7.5e-101;
Matches 442; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 GTGCCAAGCATGACTATACCTACCAAAACAAGTAGCCGCTCGGTTCTT 60
DB 182 GTGCCAAGCATGACTATACCTACCAAAACAAGTAGCCGCTCGGTTCTT 241
QY 61 CTCGTCCTCTCTCTGTGCTGGCGAGGAGCATCTTCAAGCTGTATGAGAAATTCCT 120
DB 242 CTCGTCCTCTCTCTGTGCTGGCGAGGAGCATCTTCAAGCTGTATGAGAAATTCCT 301
QY 121 TGTCTTCATATCTCTCTATTCATTCATCGTGGAGATCTACAGAAATGTTCTTCTGAGTGA 180
DB 302 TGTCTTCATATCTCTCTATTCATTCATCGTGGAGATCTACAGAAATGTTCTTCTGAGTGA 361
QY 181 TCAGCAGCTGTTGTTGAGAGCTGGCTCTGTACTGCAGACACTACATTGAGCTCATCCC 240
DB 362 TCAGCAGCAGCAGTTTGAAGAGTTGGCTCTGTACTGCAGACACTACATTGAGCTCATCCC 421
QY 241 TATATCTCTGTTCTGGGTTTCTATGTTACATTTGGTGTGAGCCGCTGGTGGACCAATA 300
DB 422 TATATCTCTGTTCTGGGTTTCTATGTTACATTTGGTGTGAGCCGATGGTGAACCAATA 481
QY 301 CGAAGACTTGGCGTGGCCGACCGCTCATGATCCAGGTGTCATGCTTCGAGAGGGCAA 360
DB 482 CGAAGACTTGGCGTGGCCGACCGCTCATGATCCAGGTGTCATGCTTCGAGAGGGCAA 541
QY 361 GGAATGAGAGAGCCGTTTGTCTGGCGGCGACGCTCATCCGTAACCCCTGGGCAAGT 420
DB 542 GGAAGAGAGAGCCGTTTGTCTGGCGGCGACGCTCATCCGTAACCCCTGGGCAAGT 601
QY 421 GCTCATCTCGGCGAGCATCAGCACTGGTTATACAGCGCTTTTC 465
DB 602 GCTCATCTCGGCGAGCATCAGCACTGGTTATACAGCGCTTTTC 646

RESULT 12
BI756228 735 bp mRNA linear EST 25-SEP-2001
LOCUS BI756228
DEFINITION 603024265F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5194649 5',
mRNA sequence.
ACCESSION BI756228

VERSION BI756228.1 GI:15747806

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Chordata; Catarrhini; Homnidae; Homo.
1 (bases 1 to 735)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: cgabs-rc@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: L1AM1487 row: a column: 18

High quality sequence stop: 577.
Location/Qualifiers

FEATURES
source

1..735
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5194649"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/note="Organ: Brain; Vector: PCMV-SPORE; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains; age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 Kb, insert size
range 1-3 Kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 20.7%; Score 397.4; DB 4; Length 735;
Best Local Similarity 83.6%; Pred. No. 8.9e-93;
Matches 474; Conservative 0; Mismatches 91; Indels 2; Gaps 2;

QY 144 CCATCCGGAACCTCTACAGAAATGTTCTCTCCAGATGATGAGAGCGTTGTTGAGAGC 203
DB 40 CCACACTGGAACCCACCTGCTGGCCCTTACGAAAGACAGCTGATGTTGAAAC 99
QY 204 TGGCTCTGTACTGCGACAGCTACATTCACTCATCTATATCTTCTGTTGGTTTCT 263
DB 100 TGACTCTGTAATTGGCGACAGCTACATCCAGCTCATCCCATTTCTTCGTTGGCTTCT 159
QY 264 ATGTTACATTTGGTGTGAGCCGCTGTGAGGCACTATGAGAACTTGGCCGCGAC 323
DB 160 AGTGAAGCTGTGCTGAGACCGCTGTGAAACAGTACGAAACCTGCGGCGCCGAC 219
QY 324 GCTCATGATCCAGGTGTCTAGCTTCTGAGAGGCAAGATGAGGAAGCCGTTTGTGCG 383
DB 220 GCTCATGAGCCTGTGTTGCTGGCTTCTGTAAGGCAAGAGAGAGAGCCGCTGCTGC 279
QY 384 GGGCGACGCTCATCCGCTACGCGCATCTGGGCGCAAGTCTCATCTCTGGCAGCATCACA 443
DB 280 GGGCGACGCTCATCCGCTACGCGCAAGTGGGCAAGTCTCATCTCTGGCAGCATCACA 339
QY 444 CCTCGGTACAAAGCGTTTCCCACTCTTCCACCTGTGCTGAGAGTTTATGAGCC 503
DB 340 CCGCAGTCTTCAAGAGGCTTTCCCAAGCGCCGACAGCACTGTGTCAAGCAGGCTTATGACTC 399
QY 504 ATGGGAAACATTAAGCAGTTTGAGAAAGTTGGCTTACCAACAACAATTTCTGGTGCCT 563
DB 400 CCGCAGAACACAAAGCAGTTTGAGAAAGTGAAGCTTACCAACAACAATTTCTGGTGCCT 459
QY 564 GGGTGTGTTTGGCAACTTGTCAATGAAGGCTTATTTTGAAGTGAATCCGGGACACCG 623

Db 460 GGGTGTGGTTGGCAACCTGTCATGAGAGCGGTGGCTTGGAGTGATCCGGAGCCCTTA 519
Qy 624 TCCCTGCTCCAGAGCTGATGATGAGTGCTGTA-CTTGGCTACTCACTGAGACAGCTG 682
Db 520 TCCCTGCTCCAGAGCTGCTGTAACGAGATGAACACCTTGCGTTCGGTGTGAGACACTG 579
Qy 683 TATGCGCT-ACGACTGATGAATATCC 708
Db 580 TATGCTTAACGACTGATGATATCC 606

RESULT 13
BI480798 666 bp mRNA linear EST 28-FEB-2002
LOCUS H2PPE-0430 Human Retinal Pigment Epithelium (2) Homo sapiens cDNA
DEFINITION 5' similar to vitellogenin macular dystrophy (best disease, mRNA
sequence.
BI480798
VERSION BI480798.1 GI:18998607
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 666)
AUTHORS Buraczynska,M., Mears,A.J., Zarepari,S., Farjo,R., Filipova,E.,
Yuan,Y., Macnee,S.P., Hughes,B. and Swaroop,A.
TITLE Towards an expression profile of native human retinal pigment
epithelium: Identification of a non-redundant set of more than 1100
genes
JOURNAL Unpublished (2001)
COMMENT Contact: Swaroop, A.
Department of Ophthalmology and Visual Sciences
Kelllogg Eye Center, University of Michigan
540 KEC, 1000 Wall St., Ann Arbor, MI 48105, USA
Tel: 734 615 2246
Fax: 734 647 0228
Email: swaroop@umich.edu
PCR Primers
FORWARD: M13/PUC-Reverse - cccagtcacgacgtcgttaaacg
BACKWARD: M13/PUC-Forward - agcgagatacaatcctcacacg
Seq primer: M13/PUC-Reverse.
FEATURES
SOURCE Location/Qualifiers
1..666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Native Retinal Pigment Epithelium sheets"
/dev_stage="juvenile"
/clone_lib="Human Retinal Pigment Epithelium (2)"
/note="Organ: Retina; Vector: pSPORT1"

ORIGIN
Query Match 20.4%; Score 391.2; DB 4; Length 666;
Best Local Similarity 81.4%; Pred. No. 3.7e-91;
Matches 464; Conservative 0; Mismatches 105; Indels 1; Gaps 1;

Qy 2 TGGCAAGCATGACTATCACTACACAAAGAGAGCAATGCGCGCTCGCTTCTTC 61
Db 98 TGGCGGCGCATGACATCACTTACCAAGCCAGTGGTAATGCCGCTTAGGCTCCCTTC 157
Qy 62 TCGTCCCTCTCTGCTGCTGGCGAGGAGCATCTTCAAGCTGCTGTATGAGAAATTCCTT 121
Db 158 TCCCGCCCTGCTGCTGCTGGCGGAGCATCTTCAAGCTGCTGTATGAGAGTTCTTA 217
Qy 122 GTCTTCATATTCCTTACTATTCATCCGTGAGCTCTCAGAGATGTTCTTCGAGAT 181
Db 218 ATCTTCCTGCTCTGCTACTATCATCTCCGCTTATTTATAGGCTGCGCTTACGAGAA 277
Qy 182 CAGCAGCTGTGTGTGAGAGAGCTGCTGTGATCGACAGCTACATTCAGCTCATCCCT 241
Db 278 CACAGCTGATGTGTGAGAACTGACTCTGTATTCGACAGCTACATCCAGCTATCCCC 337

Qy 242 ATATCTTCGCTTCTGAGTTTCTATGTTACATTTGTTGAGACCCGCTGGTGAAGCCAGTAC 301
Db 338 ATTTCCTTCGCTGCTGGGCTTCTACGTAGCGTGGTCTGAGACCCGCTGGTGAAGCCAGTAC 397
Qy 302 GAGAACTTGGCCGCGCCGAGCCGCTCATATGATTCAGAGTGTCTGAGTTCTGAGAGGCAAG 361
Db 398 GAGAACTTGGCCGCGCCGAGCCGCTCATATGAGCTGTGCTGGGCTTCTGAGAGGCAAG 457
Qy 362 GATGAGAGAGCCGCTTTCGCGGCGACGCTCATTCGCTTACGCAATTCCTGGGCCAATG 421
Db 458 GACAGAGAGCCGCTGCTGCGGCGACGCTCATTCGCTTACGCAATTCCTGGGCCAATG 517
Qy 422 CTCATCTGCGCAGCATCAGACCTCGGTCTTACAGAGGCTTCCACTTTCACCACTG 481
Db 518 CTCATCTGCGCAGCATCAGACCTCGGTCTTACAGAGGCTTCCACTTTCACCACTG 576
Qy 482 GTGCTAGCAGGTTTATGACCCATGCGGAGACATTAAGCATTTGACAGAACTTGGGCTTACCA 541
Db 577 GTGCAAGAGGCTTTATGACTTACGAGACAGAACTGAGAACTGAGAGGCTTACCA 636
Qy 542 CACAACATTTCTGCGCTGCGCTGCGGTG 571
Db 637 CACAACATTTCTGCGCTGCGCTGCGGTG 666

RESULT 14
CA389968 639 bp mRNA linear EST 06-NOV-2002
LOCUS ca389968.y1 Human Retinal Pigment Epithelium/choroid cDNA
DEFINITION (Un-normalized, unamplified): cs Homo sapiens cDNA clone ca389968
5', mRNA sequence.
CA389968
VERSION CA389968.1 GI:24720628
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 639)
AUTHORS Wistow,G., Bernstein,S.L., Wyatt,M.K., Pariss,R.N., Behal,A.,
Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
TITLE Expressed sequence tag analysis of human RPE/choroid for the
NEIBank Project: Over 6000 non-redundant transcripts, novel genes
and splice variants
JOURNAL Wistow G
MEDLINE Section on Molecular Structure and Function
PUBMED National Eye Institute
12107410 6/331, NIH, Bethesda, MD 20892-2740, USA
22103460
Contact: Wistow G
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 104 row: h column: 03
Seq primer: M13Rpl reverse primer (ABI).
FEATURES
SOURCE Location/Qualifiers
1..639
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="ca389968"
/tissue_type="RPE/choroid"
/dev_stage="Adult"
/lab_host="EMD10B"
/clone_lib="Human Retinal pigment epithelium/choroid cDNA
(un-normalized, unamplified): cs"
/note="Organ: Eye; Vector: pCMVSPORT1; Two different donor
eyes (75-80 years old) yielded approximately 600 mg of
dissected RPE/choroid tissue. This in turn yielded 340 ug
of total RNA and 7 ug of mRNA. A directionally cloned cDNA
library in the pCMVSPORT1 vector was constructed at Life

Technologies (Rockville, MD; now part of Invitrogen Corp), essentially following the protocols of the SuperScript Plasmid System (Invitrogen Corp).
<http://www.invitrogen.com/> The library code designation was cs. For this library, cDNA inserts were cloned into the NotI/MluI sites of the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

Query Match 20.3%; Score 389.2; DB 6; Length 639;
Best Local Similarity 81.3%; Pred. No. 1.2e-90;
Matches 451; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

2 TGCACAGCAGTACTATCATCACTACAAACAAAGTAGCCAAATGCCGCTCGTTCGTC 61
85 TCCCTGGCAGTACATCATCTTACACAAAGAGTGTATGCGCTTACGCTTC 144
62 TCGTCCCTCTCTCTGCTGCGAGGAGCATCTAACAGCTGTATGAGAAATTCCT 121
145 TCCCGCTGCTGCTGCTGCGGAGCATCTCAAGCTGTATGCGAGTTCTTA 204
122 GTCTTCAATTCCTTACTATTCATCCGTGAGCTCTACAGAAATGTTCTCCAGTAT 181
205 ATCTTCTGCTCTGCTACTACATCATCCGCTTTATTTAGCTGCCCTCAGGAA 264
182 CAGCAGCTGTGTTGAGAAAGTGTGCTGTACTGCGACAGTACATTCAGTCCCT 241
265 CAACAGCTGATGTTTGAAGAACTGACTGTATTCGACAGTACATTCAGTCCCT 324
242 ATATCTCTGTTCTGCTGCTTCTATGTTAATTTGTTGAGTCCGCTGTGAGCA 301
325 ATTTCTCTGCTGCTGCTGCTTCTACGTCGTCGTCGTCGTCGTCGTCGTCG 384
302 GAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
385 GAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 444
362 GATGAGAAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
445 GACGAGCAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504
422 CTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
505 CTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564
482 GTGCTAGAGGTTTATATACCCATGAGGAGAAATAGAGTGTGAGAGTGTGAG 541
565 GTGCAAGCAGGCTATGATCTCCGCAAGCAAGAGTGTGAGAACTGAGCTTACA 624
542 CACAACATCTCTG 556
625 CACAACATCTCTG 639

RESULT 15
CV024947
LOCUS 2492 Full length cDNA from the Mammalian Gene Collection Homo
DEFINITION sapiens cDNA 5' similar to BC015220, mRNA sequence.
ACCESSION CV024947
VERSION CV024947.1 GI:51482835
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 491)
Raul, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S.,
Driscoll, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O.,
Clingsingmish, T.R., Hartley, J.L., Bessito, D., Cheo, D., Moore, T.,
Simmons, B., Segueria, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C.,
Vandenhaute, J., Cusick, M.E., Albalade, J.S., Hill, D.E. and Vidal, M.

TITLE Human ORFome Version 1.1: a Platform for Reverse Proteomics
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@fci.harvard.edu

ORF Sequence Tag (OST) of Gateway construct. Each cloned ORF
results from a PCR reaction using an MGC full-length cDNA as
template DNA and ORF specific primers
PCR Primers
FORWARD: ATGTTGAGAACTGACTCTGAT
BACKWARD: TAGGAATGCTTCACTCCGT
Insert Length: 491 Std Error: 39.00
Plate: 11066 row: 05 column: G
Seq primer: ACTGCGCTGCTGTTTACAGCTGCTGAGTGGGAAAC
High quality sequence start: 94
High quality sequence stop: 490
POLYA=No.

FEATURES
source
location/Qualifiers
1..491
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="mixed"
/clone_lib="Full length cDNA from the Mammalian Gene
Collection"

/note="Vector: mixed; The ORFs were PCR amplified from the
MGC (Mammalian Gene Collection) as of April 2004 and
cloned by recombinational Gateway cloning into pDONR223
Donor vector. Reference : MGC (Mammalian Gene Collection)
Program Team. Generation and Initial Analysis of more than
15,000 Full-length Human and Mouse cDNA Sequences. PNAS,
2002, 99(26), 16899-16903"

ORIGIN
Query Match 19.9%; Score 381.2; DB 7; Length 491;
Best Local Similarity 86.1%; Pred. No. 1.5e-88;
Matches 422; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

192 TGTTTGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 251
2 TGTTTGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 61
252 TTCTGAGTTTCTATGTTTACATTTGCTGAGCCGCTGCTGCTGCTGCTGCTG 311
62 TGCTGAGCTTCTACGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 121
312 CGTGGCCGAGCGCTCATGATCCAGGTGTCTAGCTTGTGAGAGGCAAGATGAG 371
122 CGTGGCCGAGCGCTCATGATCCAGGTGTCTAGCTTGTGAGAGGCAAGATGAG 181
372 GCCGTTGCTGCGGCGACGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 431
182 GCCGCTGCTGCGGCGACGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
432 GCAGCATCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 491
242 GCAAGCTCAGACCGAGTCTTACAAAGCTTCCCGAGCGCCAGCACTGTGTGAG 301
492 GTTTTATGACCAATGAGGAAATATAGAGTGTGAGAGTGTGAGAGTGTGAGAG 551
302 GCTTTATGACTCCGCGAGAAACAGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 361
552 TCTGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 611
362 TCTGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
612 TCCGAGACACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 671
422 TCCGAGACCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481

Qy	672 GTGGACACT	681
Db	482 GTGGACACT	491

Search completed: March 26, 2005, 13:32:12
Job time : 6165.57 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 21:57:50 ; Search time 993.266 Seconds
(without alignments)
11419.125 Million cell updates/sec

Title: US-09-622-964A-28

Perfect score: 1916
Sequence: 1 gtcgcaagcatgactatca.....aaaaaaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseq1980s:*
2: geneseq1990s:*
3: geneseq2000s:*
4: geneseq2001s:*
5: geneseq2001as:*
6: geneseq2002as:*
7: geneseq2002bs:*
8: geneseq2003as:*
9: geneseq2003bs:*
10: geneseq2003cs:*
11: geneseq2003ds:*
12: geneseq2004as:*
13: geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	# Match	Query Length	DB ID	Description
1	1916	100.0	1916	2	AAZ21229
2	871.2	45.5	2229	2	AAZ21227
3	867	45.3	1758	8	ABZ80972
4	658.2	34.4	2404	12	ADQ84435
5	658.2	34.4	2404	13	ADQ83262
6	658.2	34.4	2429	2	AAZ21228
7	519.4	27.1	2112	6	ABQ61177
8	519.4	27.1	2137	8	ABZ80973
9	519	27.1	1530	8	ABZ80997
10	458.2	23.9	2500	11	ADM02614
11	423	22.1	2028	8	ABZ80974
12	401.2	20.9	1948	9	ACC59900
13	399.2	20.8	1422	8	ABZ80975
14	296.2	15.5	2861	4	ABL10793
15	260.4	13.6	1526	5	AAZ72796
16	229	12.0	1608	4	ABL12609
17	227.6	11.9	1263	2	AAV99722
18	215.4	11.2	10760	4	ABL10792
19	211.2	11.0	1592	10	ADC29982
20	194.4	10.1	620	5	ADL45614

21	194.4	10.1	1345	4	ABL12637
22	190.8	10.0	305	5	ABAI1094
23	188.4	9.8	7108	5	ABAI4556
24	188.4	9.8	16125	2	AAZ21226
25	188.4	9.8	16650	5	ABAI4559
26	188.4	9.8	18530	5	ABAI4557
27	188.4	9.8	18537	5	ABAI4558
28	187.8	9.8	1350	4	AAZ76848
29	187.2	9.8	2196	4	ABL16195
30	181.2	9.5	1292	11	ACN89920
31	166.6	8.7	3592	4	ABL12636
32	163	8.5	1238	9	ADA44960
33	129.6	6.8	3785	11	ACN90196
34	122.2	6.4	539	12	ACH74388
35	119.8	6.3	235	12	ACH88088
36	118	6.2	2914	12	ADQ24621
37	118	6.2	2914	12	ADQ25483
38	114.4	6.0	853	11	ACN90196
39	113.8	5.9	1815	10	ADF82072
40	113.8	5.9	1954	12	ADQ24756
41	113.4	5.9	1330	10	ADC31917
42	105.6	5.5	2104	5	AAZ72795
43	95.4	5.0	3592	4	ABL12636
44	95	5.0	527	4	ABL17589
45	94.4	4.9	392	4	ABL11210

ALIGNMENTS

RESULT 1
AAZ21229
ID AAZ21229 standard; CDNA; 1916 BP.
XX
AC AAZ21229;
XX
DT 22-NOV-1999 (first entry)
XX
DE Mouse CGICE CDNA sequence.
XX
KW CGICE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
RV age-related macular dystrophy; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 11..166
FT FT /*tag= a
FT FT /product= "CGICE protein"
XX
PN W09943695-A1.
XX
PD 02-SEP-1999.
XX
PF 22-FEB-1999; 99WO-US003790.
XX
PR 25-FEB-1998; 98US-0075941P.
PR 18-DEC-1998; 98US-0112926P.
XX
PA (MERI) MERCK & CO INC.
PA (UYU-) UNIV UPPSALA.
XX
PI Petrunkin K, Caskey CT, Metzker M, Wadelius C;
XX
DR WP1; 1999-540560/45.
DR P-PSDB; AAY29955.
XX
PT Human and mouse polynucleotides encoding CGICE polypeptides.
XX
PS Claim 2; Fig 8; 67pp; English.
XX
CC The present sequence represents the mouse CGICE CDNA sequence, which when
mutated is responsible for Best's macular dystrophy (BMD).

CC Polynucleotides encoding CGICE are useful for diagnosing whether a
CC patient carries a mutation in the CGICE gene. Normal and mutated CGICE
CC proteins are useful for identifying activators and/or inhibitors of these
CC proteins, in order to treat BMD. The CGICE gene offers a simpler and
CC cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
CC age-related macular dystrophy
XX

Sequence 1916 BP; 514 A; 499 C; 457 G; 446 T; 0 U; 0 Other;

Query Match 100.0%; Score 1916; DB 2; Length 1916;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1916; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GTGCCAAGCCATGATCTATCATCAACAACAAAGTACCAATGCGCCGCTCGTTCCT 60
1 GTGCCAAGCCATGATCTATCATCAACAACAAAGTACCAATGCGCCGCTCGTTCCT 60
61 CTGCTCCCTCTCTGCTGCTGCGAGGAGCATCTACAAAGCTGTGTATGAGAAATTCCT 120
61 CTGCTCCCTCTCTGCTGCTGCGAGGAGCATCTACAAAGCTGTGTATGAGAAATTCCT 120
121 TGTCTCAATATCCCTCTACTATTCATCCGCTGAGCTCTACAGAAATGTTCTCTGAGTGA 180
121 TGTCTCAATATCCCTCTACTATTCATCCGCTGAGCTCTACAGAAATGTTCTCTGAGTGA 180
121 TGTCTCAATATCCCTCTACTATTCATCCGCTGAGCTCTACAGAAATGTTCTCTGAGTGA 180
181 TCAGCAGCTGTGTGTTGAGAAAGCTGCTCTGACTGCGACAGCTCATTCAGCTCATCC 240
181 TCAGCAGCTGTGTGTTGAGAAAGCTGCTCTGACTGCGACAGCTCATTCAGCTCATCC 240
241 TATATCTCTCGTTCGTGGGTTTCTATGTTACATTTGGTGTGAGCCGCTGTGAGCACTGA 300
241 TATATCTCTCGTTCGTGGGTTTCTATGTTACATTTGGTGTGAGCCGCTGTGAGCACTGA 300
301 CGAGAAGCTTGCCTGCGGCGCGACCGGCTGATATCCAGAGTGTCTAGCTTCGTGAGAGGCA 360
301 CGAGAAGCTTGCCTGCGGCGCGACCGGCTGATATCCAGAGTGTCTAGCTTCGTGAGAGGCA 360
361 GGATGAGAGAGGCGCTTGTGTGCGGCGCGACCGCTCATCCGCTACGCGCATCTGCGGCAAGT 420
361 GGATGAGAGAGGCGCTTGTGTGCGGCGCGACCGCTCATCCGCTACGCGCATCTGCGGCAAGT 420
421 GCTCATCTCTGCGAGCATCAGACCTCGGTCTACAGCGCTTCCACTCTTCAACCACT 480
421 GCTCATCTCTGCGAGCATCAGACCTCGGTCTACAGCGCTTCCACTCTTCAACCACT 480
481 GGTGTACAGAGTTTATGACCAATGGGGAACATTAAGAGTTGAGAAAGTGGGCTTACC 540
481 GGTGTACAGAGTTTATGACCAATGGGGAACATTAAGAGTTGAGAAAGTGGGCTTACC 540
541 ACACAAACATTTCTGGGTCCTGCGGTGTGTTGCCAATTTGTCATTAAGAGGCTTACT 600
541 ACACAAACATTTCTGGGTCCTGCGGTGTGTTGCCAATTTGTCATTAAGAGGCTTACT 600
601 TGGAGTCTGAGTCCGGGACACCGTCTGCTCAGAGGCTGTGATAGGGGTGTACTTT 660
601 TGGAGTCTGAGTCCGGGACACCGTCTGCTCAGAGGCTGTGATAGGGGTGTACTTT 660
661 GCGTACTCAGTGTGAGAGCTGTATGCTTACGACTGATAGTATCCATTGGTGTGAC 720
661 GCGTACTCAGTGTGAGAGCTGTATGCTTACGACTGATAGTATCCATTGGTGTGAC 720
721 ACAGAGTGTGACAGTGTGAGTATACAGTTTCTTTCATGCTGTGATCGGAGGAGGTT 780
721 ACAGAGTGTGACAGTGTGAGTATACAGTTTCTTTCATGCTGTGATCGGAGGAGGTT 780
781 TGTGAACCCCAACAAAGACTACCCAGAGCAGATGATCTGTGTTGTGCTGTCTTAC 840
781 TGTGAACCCCAACAAAGACTACCCAGAGCAGATGATCTGTGTTGTGCTGTCTTAC 840
841 AATCTGCAATTTCTATTTACATGGGCTGGCTGAAGGTGGCAGAAACAGCTCATCAACC 900
841 AATCTGCAATTTCTATTTACATGGGCTGGCTGAAGGTGGCAGAAACAGCTCATCAACC 900

901 CTTGCGGAGAGCAGATGATGATTTTGGACCTAATCTGATCATTTGACAGAAACCTGACGT 960
901 CTTGCGGAGAGCAGATGATGATTTTGGACCTAATCTGATCATTTGACAGAAACCTGACGT 960
961 GTCCCTGTGTGCTCGTGTGATGAGATGACACAGAACTTGCCTCCATGAGAAAGTGTACATGTA 1020
961 GTCCCTGTGTGCTCGTGTGATGAGATGACACAGAACTTGCCTCCATGAGAAAGTGTACATGTA 1020
1021 CTGGAAGAGGACGAGCGCTTACGCGCCCTTACACAGCTGTCTTCTGCGAGTCTCGCGGCA 1080
1021 CTGGAAGAGGACGAGCGCTTACGCGCCCTTACACAGCTGTCTTCTGCGAGTCTCGCGGCA 1080
1081 TTCCCTATGAGGCTCCACTTTCACATCAGCTTAAAGAAAGAACTTAAAGCTTTGCTG 1140
1081 TTCCCTATGAGGCTCCACTTTCACATCAGCTTAAAGAAAGAACTTAAAGCTTTGCTG 1140
1141 AAAAGAGAGGCTGACACGAGATTAAGAAAGAGTGTATAGACACCATAGGCTGCTT 1200
1141 AAAAGAGAGGCTGACACGAGATTAAGAAAGAGTGTATAGACACCATAGGCTGCTT 1200
1201 CTTAGGACTGACACCCCAAACTTACATCTTCCCTTGAAGAACTTAAAGCCTTATTT 1260
1201 CTTAGGACTGACACCCCAAACTTACATCTTCCCTTGAAGAACTTAAAGCCTTATTT 1260
1261 GTGTTTAAAGAACCCCTCTCTGAAAGGCGAGTGAAGAGGCCAACAGAAAAACAGAA 1320
1261 GTGTTTAAAGAACCCCTCTCTGAAAGGCGAGTGAAGAGGCCAACAGAAAAACAGAA 1320
1321 AGATGTCTGAAATTTAAGGCTGTGACTTCTTGAATGTGTTCAAGGTTTAAAGAGAG 1380
1321 AGATGTCTGAAATTTAAGGCTGTGACTTCTTGAATGTGTTCAAGGTTTAAAGAGAG 1380
1381 AGGTCTCCATTTGTGGCCACAGGACCCAGACACCCCTTACGAGACAGTCAACACCTTC 1440
1381 AGGTCTCCATTTGTGGCCACAGGACCCAGACACCCCTTACGAGACAGTCAACACCTTC 1440
1441 CAGTTCAAGACAGAGTGTAGGCGCTTCAACAGATTTACAAAGAAATCTGTCACTGAAAAA 1500
1441 CAGTTCAAGACAGAGTGTAGGCGCTTCAACAGATTTACAAAGAAATCTGTCACTGAAAAA 1500
1501 GAAAACTGTGTGAGTTTAACTTGAACATTCAGAGAGCCCAAGAACTTCTTCAACAGCG 1560
1501 GAAAACTGTGTGAGTTTAACTTGAACATTCAGAGAGCCCAAGAACTTCTTCAACAGCG 1560
1561 CCGTTTGAACAGATGTCAACCAATTAACAGGCTTAAAGAGAGCAGTGCAGTCTTA 1620
1561 CCGTTTGAACAGATGTCAACCAATTAACAGGCTTAAAGAGAGCAGTGCAGTCTTA 1620
1621 TCCCTACAGGATGAGCTGTGACCAAACTGTTCTTATGAGTGTGCTCAGACCTG 1680
1621 TCCCTACAGGATGAGCTGTGACCAAACTGTTCTTATGAGTGTGCTCAGACCTG 1680
1681 GCCCTGACTTGCAGAGATGCCAGCGGCGACTGACCCAGTCAAAAGGCAACAGAGCG 1740
1681 GCCCTGACTTGCAGAGATGCCAGCGGCGACTGACCCAGTCAAAAGGCAACAGAGCG 1740
1741 ACACCCAGAGAGTGTGTTCCACAGACAGTCAAGATGTAAGTCAAGAACTTAA 1800
1741 ACACCCAGAGAGTGTGTTCCACAGACAGTCAAGATGTAAGTCAAGAACTTAA 1800
1801 TAGTCTGCTGAAAAACCTGTATTTTACATCTTCCCAAACTTAAGAGTTTAAATAA 1860
1801 TAGTCTGCTGAAAAACCTGTATTTTACATCTTCCCAAACTTAAGAGTTTAAATAA 1860
1861 CCGTAATATTTCTTTTGTGAGAAAAAATTTTGTGAGAAAAAATTTTGTGAGAAAAA 1916
1861 CCGTAATATTTCTTTTGTGAGAAAAAATTTTGTGAGAAAAAATTTTGTGAGAAAAA 1916

RESULT 2
AAZ21227
ID AAZ21227 standard; cDNA; 2229 BP.

QY 1355 AATGTGTTCCAGGATTTAAGAGAGAGGCTCCCATTTGTGGCCCAAGGACCCAGCAGC 1414
DB 1464 AAGCTGCCCCACCTGTATACAGAGGCGAGCTACTACAGTCCCAAGAGGCCCTTCAGC 1523
QY 1415 CACCTTAC 1422
DB 1524 CCCACTCC 1531

RESULT 3
AB280972
ID AB280972 standard; DNA; 1758 BP.
XX
AC AB280972;
XX
DT 15-OCT-2003 (first entry)
XX
DE Human bestrophin (vitelliform macular dystrophin 2) gene.
XX
KW Bestrophin; vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;
KW antiobesity; antidiabetic; immunomodulator; hypotensive; cardiac;
KW antileptic; osteopathic; antiinflammatory; cyostatic; obesity;
KW energy homeostasis; metabolism; triglyceride; body-weight regulation;
KW eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
KW coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
KW sleep apnea; chromosome 11.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1758
FT FT /tag= a
FT FT /product= "VMD2"

XX W02003030922-A2.
XX PD 17-APR-2003.
XX PF 09-OCT-2002; 2002WO-BP011321.
XX PR 09-OCT-2001; 2001EP-00124059.
XX PA (DEVE-) DEVELOPENTWICKLUNGSHIOLOGISCHE FORSCH.
XX PI Steernagel A, Broenner G, Fritsch R, Eulenberg K, Ciosek T;
XX DR WPI; 2003-393411/37.
XX DR P-PSDB; ABR58055.
XX PT New pharmaceutical compositions comprising a Bestrophin gene, polypeptide
XX PT or nucleic acid, for treating, alleviating and/or preventing metabolic
XX PT diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
XX PT gallstones.
XX PS Claim 2; Fig 4a; 85pp; English.
XX CC This sequence represents the coding region for a member of the human
XX CC bestrophin gene family designated vitelliform macular dystrophin 2
XX CC (VMD2). The dystrophin gene family are involved in energy homeostasis and
XX CC metabolism of triglycerides. The sequence can be used for the manufacture
XX CC and/or preventing disorders including metabolic diseases such as obesity
XX CC and other body-weight regulation and related disorders such as eating
XX CC disorder, cachexia, diabetes mellitus, hypertension, coronary heart
XX CC disease, hypercholesterolemia, osteoarthritis, gallstones, cancers of the
XX CC reproductive organs, and sleep apnea. The gene is found on human
XX CC chromosome 11
XX SQ Sequence 1758 BP; 421 A; 515 C; 448 G; 374 T; 0 U; 0 Other;

Query Match 45.3%; Score 867; DB 8; Length 1758;
Best Local Similarity 77.1%; Pred. No. 1.3e-209;

Matches 1110; Conservative 0; Mismatches 290; Indels 39; Gaps 3;
QY 11 ATGACTACACCTACACAAACAAAGTAGCCAAATGCCCGCCCTGGTTGTTCTGCTCC 70
DB 1 ATGACATCATCTTACACAGCCAAAGTGTCTAATGCCCTTAGAGCTCTTCTCCCGCTG 60
QY 71 CTCTGTGCTGCGAGGAGCATCTACAGCTGTGTATGAGAAATTCCTTGTCTTATA 130
DB 61 CTGCTGTGCTGCGAGGAGCATCTACAGCTGTGTATGAGAAATTCCTTGTCTTATA 120
QY 131 TTCCTACTATTCATTCCTGAGACTCTACAGAAATGTTCTTCAGATGATGACAGCTG 190
DB 121 CTGCTACTATCATCATCCGCTTATTTATTAAGGCTGCGCTCACAGAAACAAAGCTG 180
QY 191 TTGTTGAGAAAGCTGCTGTACTGAGCAGCTACATTCAGCTCATCCCTATATCCCTTC 250
DB 181 ATGTTGAGAAAGCTGCTGTATTTGAGACGTTATTCAGCTCATCCCTATATCCCTTC 240
QY 251 GTTCTGAGTTTCTATATGTTACATTTGTGTAGCGCTGTTGAGCCAGTACGAACTTG 310
DB 241 GTGCTGAGCTTCTACAGTACGCTGTGTGTGACCCGCTGTTGAGCCAGTACGAACTTG 300
QY 311 CCGTGCCGACGCGCTCATATCCAGGTGTCTAGCTTGTGAGGCGCAAGATGAGAA 370
DB 301 CCGTGCCGACGCGCTCATATCCAGGTGTGTGTGACCCGCTGTTGAGCCAGTACGAACTTG 360
QY 371 GCGCTTGTGCTGCGGCGACGCTCATCCGCTACGCAATCTGCGCCAAAGTGTCTCATCTG 430
DB 361 GCGCGGCTGCTGCGGCGACGCTCATCCGCTACGCAATCTGCGCCAAAGTGTCTCATCTG 420
QY 431 CCGAGATCAGACCTCGGCTTACAAAGCGCTTCCCACTTTCACACACCTGCTGTACGA 490
DB 421 CCGAGCTACAGACCTCGGCTTACAAAGCGCTTCCCACTTTCACACACCTGCTGTACGA 480
QY 491 GGTTTTATGACCCATGAGGAGCAATTAAGCAAGTTGAGGCTTACCAACACAAACA 550
DB 481 GGTTTTATGACCCATGAGGAGCAATTAAGCAAGTTGAGGCTTACCAACACAAACA 540
QY 551 TTCTGAGTCCCTGAGGT 610
DB 541 TTCTGAGTCCCTGAGGT 600
QY 611 ATCCGAGACACCGCTGCTGCTGACAGCTGTATGAATGAGGTGTGTGTGTGTGTGTGT 670
DB 601 ATCCGAGACACCGCTGCTGCTGACAGCTGTGTATGAATGAGGTGTGTGTGTGTGTGT 660
QY 671 TGTGACAGCTGTATGCTTACAGCTGTATGAATGAGGTGTGTGTGTGTGTGTGTGT 730
DB 661 TGTGACAGCTGTATGCTTACAGCTGTATGAATGAGGTGTGTGTGTGTGTGTGTGT 720
QY 731 ACAGTGCAGTATACAGCTTTTCTGTCATGCTGTATCCGAGAGGAGCTTCTGACACCA 790
DB 721 ACTGTGCGGAGTATACAGCTTTTCTGTCATGCTGTATGAATGAGGTGTGTGTGTGTGT 780
QY 791 AACAGACATACCCAGGCGATGAGATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 850
DB 781 GCCAAGGCTTACCTGCGCATGAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
QY 851 TTCTTATTCATGAGGT 910
DB 841 TTCTTATTCATGAGGT 900
QY 911 GACGATGATGATTTTGAAGACTTACATGATGATGATGATGATGATGATGATGATGATG 970
DB 901 GATGATGATGATTTTGAAGACTTACATGATGATGATGATGATGATGATGATGATGATG 960
QY 971 TCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1030
DB 961 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1031 GCAGCGCTCAGCGCGCTTACAGCTGCTTCTGCAAGCTGCGCGGAGCTTCTTCTCATG 1090
DB 1021 CCGAGGCAAGCGCGCTTACAGCTGCTTCTGCAAGCTTCTGCGGAGCTTCTTCTCATG 1080

Qy	1091	GGCTCCACCTTCAACATCATGCTTAAAGAAAGAAAGACTTGAAGCTTTGGTCAAAAGAGAG	115
Db	1081	GGCTCCACCTTCAACATCATGCTTGAACAAAGAGAGATGAGATTGCCAGCTCAATCATGAG	114
Qy	1151	GCTGACACGGAATAAGAAAGAGTGGCTATTAGACAGCAACATAAGGTGCTTTCTTAGAGACTG	121
Db	1141	G-----ACGAGAGAGATGCTACGCTGGCATCTTAGCGGCTTCTTAGGCTTG	118
Qy	1211	CAACCCAAAACCTACCATCTTCTCCCTGAAAAGACTTAAAGACCAACTATTTGTGTTAAAG	127
Db	1189	CAGTCCCATGATCACCATCTCTCCAGGGGCAAACTCAAGAGCCAAACTACTGTGGCCCAAG	124
Qy	1271	A-----ACCCCTCTCTCGAAGGCGAGTGTAAAGATGCCAACAGAAAAACCGAAAG--	132
Db	1249	AGGAATCCCTTCTCCACGAGGCGCTGCCCAAAAACACAGACGACCAACGAAGCGTT	130
Qy	1323	-----ATGTCGGAATTTAAAGGCTGTGAGACTTCTTGAATGTGTT	136
Db	1309	AGGGGCGCAGAAAGACACAAAGCGCTGGAAGCTTTAGGCTGTGACCGCTTCAAGTCTGCC	136
Qy	1364	CCAAGGTTTAAAGAGAGAGGCTTCCATTGTGGCCACAGGCAACCCAGACCACTTAC	1422
Db	1369	CCACTGTATCAGAGGCGAGGCTACTACAGTGGCCCAACAGACCCCTCAGCCCACTGC	1427
RESULT 4			
ADQ84435			
ID	ADQ84435	standard; cDNA; 2404 BP.	
AC	ADQ84435;		
XX	07-OCT-2004	(first entry)	
DE	Human tumour-associated antigenic target (TAT) cDNA sequence #1249.		
XX			
KM	human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;		
KW	cancer; cell proliferative disorder; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2004060270-A2.		
PD	22-JUL-2004.		
XX			
PF	15-OCT-2003; 2003WO-US029126.		
XX			
PR	18-OCT-2002; 2002US-0418988P.		
XX			
PA	(GENTH) GENENTECH INC.		
PA	(WUTD/) WU T D.		
PA	(ZHOU/) ZHOU Y.		
XX			
PI	Wu TD, Zhou Y;		
XX			
DR	WPI; 2004-534300/51.		
XX			
PT	New nucleic acid molecule and encoded polypeptide, for diagnosing,		
PT	preventing or treating cell proliferative disorders such as cancer.		
XX			
PS	Claim 1; SEQ ID NO 1249; 5504bp; English.		
XX			
CC	The present invention describes an isolated tumour-associated antigenic		
CC	target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide		
CC	sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of		
CC	(a); (c) the complement of (a) or (b); (d) a sequence that has 80%		
CC	sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-		
CC	(c). Also described: (1) an expression vector comprising the above		
CC	nucleic acid; (2) a host cell comprising the above expression vector; (3)		
CC	a process for producing a polypeptide; (4) an isolated polypeptide		
CC	comprising: (a) an amino acid sequence encoded by any of the above		
CC	nucleotide sequences; (b) an amino acid sequence encoded by the full-		
CC	length coding region of the above nucleotide sequences; or (c) a sequence		

	having at least 80% identical to a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumor comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumor in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytotoxic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing a medicament for the therapeutic treatment or diagnostic detection of a cell proliferative disorder or cancer. The present sequence represents a human TAT cDNA sequence from the present invention.
SQ	Sequence 2404 BP; 590 A; 696 C; 611 G; 507 T; 0 U; 0 Other;
Query Match	34.4%; Score 658.2; DB 12; Length 2404;
Best Local Similarity	67.6%; Pred. No. 1.7e-156;
Matches 1116; Conservative	0; Mismatches 293; Indels 242; Gaps 4;
OY	2 TGCCAGCGCATGACTACCTACCAACAAGAAGGCATTGCCGGCTGGTTTCCTC 61
Db	96 TGCCTGGCGAAGACATCATCTTAACAACAAGAGGTAAATGCCGGTAGGCTCTTC 155
OY	62 TCGTCCCTCCCTCGTGCTGCGAGAGCGAGACTTACAAGCTGCTATAGGAATTCCCT 121
Db	156 TCCGGCCCTGCTGCTGTGCTGGCGGGGAGCATCTAACAGCTGCTATAGGGAATCTTA 215
OY	122 GTCTTCATATTCTCTACTATTTCATCCGCGGACTTACAGAAATGGTCTCTCGAGTAT 181
Db	216 ATCTTCCTGCTCTGCTACTACTATCATCCGCTTTATTATAGCGTGGCCCTCACGGAA 275
OY	182 CAGCAGCTGTGTTTGAGAACTGGCTCTGTACTGCGACAGCTACATTCACTATCCCT 241
Db	276 CAACAGCTGAATGTTGGAATACTAATCTGTATTGCGACACTCATCACCCTCAATCCC 335
OY	242 ATTATCCTTGCTTGGGGTTCATATGTTACATTTGGTGGAGCCGCTGGTGGAGCCATGAC 301
Db	336 ATTTCCTTCGTGCTGGGCTTTTACGTGACGCTGGTCTGACCCGCTGTGGAAACAATGAC 395
OY	302 GAGAACTTGCCGTGAGCCGACCGGCTCATGATTCAGAGTGTAGCTTGTGGAAGGCAAG 361
Db	396 GAGAACTGCGGTGGCCGACCGGCTCATGAGCCGAGTGGTGGGCTTGTGGAAGGCAAG 455
OY	362 GATGAGGAAGGCGCTTTGCTGCGGCGCAGCTTACCGCTACGCCATCTTGGGCAAGTG 421
Db	456 GACAGGACAGGCGCGGCTGCTGGCGCGCACGCTCATCCGCTACCGCCAACTTGGCAACGTG 515
OY	422 CTCAATCCGCGGACATCAGCACTCGGCTCAACAAGGCTTTCACACTTACACACTG 481
Db	516 CTCAATCCGCGGACGCTCAGCACTCGGCTCAACAAGGCTTTCACAGCGCCACAGCACTG 575
OY	482 GTGCTAGCAGGTTTTATGACCATGAGGGAACATAAGCAGTTGAGAAATTGGGCTTACCA 541
Db	576 GTGGAAGAGGCTTTATGACCTCGGCAAGAACACAAGCAGTTGGAAGAACTGAGCTACCA 635
OY	542 CACAACACATTTCTGGGTGCTCGGTGTGGTTTGCCAACTTGTCAATGAAAGCCATATTT 601
Db	636 CACAACATGTTTCTGGGTGCTCGGTGTGGTTTGCCAACTTGTCAATGAAAGCCATATTT 695

QY	602	GGAGGTGGAAATCCGGGGACACCGTCCTCTCAAGCTCATGAAATGAGGTGTGTACTTTG	661
Db	696	GGAGGTGGAATCCGGGGACCCCTATCTCTCCAGAGCCTGTCTGAACAGATGAACACCTTG	755
QY	662	CGTACTCAGTGTGGACAGCTGTATGTGCCTACGACGTGATAGTATCCCATTTGGGTACACA	721
Db	756	CGTACTCAGTGTGGACACTGTATAGCTTACGATCGATTAAGTATCCCATCTGGTGTATACA	815
QY	722	CAGTGTGTGACAGTGGCAGATATACAGCTTTTCTCTTGATGCTTGTGATCGGAGGACGTTT	781
Db	816	CAGTGTGTGCTGTGGGTGGTGAACAGCTTCTCTCACTGTCTTAAGTGTGGGGCGGACGTTT	875
QY	782	CTGAACCCAAACAAGACTACCCAGGCGCATGATGATGTGGTGTGGCTGTCTTCAACA	841
Db	876	CTGAACCCAGCCAAAGGCTTACCTCTGGCCATGAGCTGAACTCGTTGTGCCCGCTTCACG	935
QY	842	ATCTCTCAATCTTTATTTCTATCAATGSGCTGGCTGAA-----	876
Db	936	TTCTCTCAGATTCTTCTTATATGTGGCTGGCTGAAGTGGGCCTTCCAGGGCCCGTGG	995
QY	877	-----	876
Db	996	GGCTGAGGCAATGGCCAGAGGGGTCAATGCCACAGCACTGCTTGAGACGAGATGCATGT	1055
QY	877	-----	876
Db	1056	CAGAAAGGAAGGTCTCAGGGGTAGAAACACACGACGCGCTGTGGCGCACACTGTAAATC	1115
QY	877	-----	878
Db	1116	CCAGCTACTCGGGAGGCTGAGGCAAGAAATCGCTTGAACCCGGGAGCGAGGTGTGTGG	1175
QY	879	TGGCAGAACAGCTCATCAACCCCTTCGSGGAGAGACGATGATTTTGACCTAACCTGGA	938
Db	1176	TGGCAGAGCAGCTCATCAACCCCTTTGGAGAGAGATGATATTTTGAGACCAATCTGGA	1235
QY	939	TCATTGACAGAAACCTGCAGGTGTCCCTGTTGTCCGTGATGGATGACACAGAACTTGC	998
Db	1236	TTGTGTGACAGGAATTTGACAGGTGTCCCTGTTGTGCTGTGGATGAGATGACACGAGGACCTGC	1295
QY	999	CTCCATGGAAAGTGCATATGTACTGGAACGAGGCGAGCCCTCAGCGGCCCTTACACAGCTG	1058
Db	1296	CTCGATGAGACCGGACATGATCTGGAATAGCCCGAGACCAAGCCCTTACACAGCTG	1355
QY	1059	CTTTCGCAAGCTTGCAGCGGCAATTCCTTCATGAGGCTCCACCTTCAACATCAGCCTTAAGA	1118
Db	1356	CTTCGCGCAAGTTCGTCGAGGCTCTTTATATGGGCTCCACTTCAACATCAGCCTTAAGA	1415
QY	1119	AAGAAGACTTGAAGCTTTGTGTCMAAAGAGAGGCTGACACGATTAAGAAAGAGTGGCT	1178
Db	1416	AAGAAGAAATGAGATTCCAGGCCAATCGAGG-----ACAGAGGAGATGCTC	1463
QY	1179	ATAGCAGCACATAGGCTGTCTTCTTAGGACTGCACCAACCAAAATACCATCTTCCCTTGA	1238
Db	1464	AAGCTGGCATATATGGCCGCTTCTTAGGCTGTCAAGTCCCATGATCACATCTTCCACGGG	1523
QY	1239	AAGACTTAAAGCAAAACTATTTGTGTCTTAAGAA-----CCCGCTCTCGAAGGCGAGT	1292
Db	1524	CAAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCTTCCAGAGGGGCTTGC	1583
QY	1293	GTAAGGAATGCCAACCAAGAAAACCAAGAAAGT-----GTCTGGA	1331
Db	1584	CCAAAACCAACAAAGCACCCAACGAACGTTAAGGGCCAGGAACACAAAGCCCTGGA	1643
QY	1332	AATTTAAAGGCTTGTGACTTCTTGAATATGTGTTCCAAAGTTTAAAGAGAGGCTCCCATT	1391
Db	1644	AGCTTAAAGGCTGTGAGCGCCCTTCAAGTCTGGCCCACTGTATCAGAGGCCAGGCTTACTACA	1703
QY	1392	GTGGCCCAACAGGCAACCAAGACGCAACCTTAC	1422
Db	1704	GTGGCCCAACAGAGCGCCCTTACAGCCCACTGC	1734

RESULT 5
 ID ADO83262 standard; cDNA; 2404 BP.
 AC ADO83262;
 DT 07-OCT-2004 (first entry)
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #76.
 KM human; tumour-associated antigenic target; TAT; cytostatic; gene therapy
 KW cancer; cell proliferative disorder; gene; sa.
 OS Homo sapiens.
 PN WO2004060270-A2.
 XX 22-JUL-2004.
 PF 15-OCT-2003; 2003WO-US029126.
 PR 18-OCT-2002; 2002US-0418988P.
 PA (GENTH) GENENTECH INC.
 PA (MUTD) MU T D.
 PA (ZHOV) ZHOU Y.
 PI Mu TD, Zhou Y;
 XX WPI; 2004-534300/51.
 DR New nucleic acid molecule and encoded polypeptide, for diagnosing,
 XX preventing or treating cell proliferative disorders such as cancer.
 PT
 PS Claim 1; SEQ ID NO 76; 5504dp; English.
 XX
 XX The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO.1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a) -
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector;
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a

CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAR CDNA sequence from the present invention.
 XX

Sequence 2404 BP; 590 A; 696 C; 611 G; 507 T; 0 U; 0 Other;

Query Match 34.4%; Score 658.2; DB 13; Length 2404;
 Best Local Similarity 67.6%; Pred. No. 1.7e-156;
 Matches 116; Conservative 0; Mismatches 293; Indels 242; Gaps 4;

```

QY 2 TGGCAAGCATGATCATCTACCTACCAACAAAGTACCAATGCCCCGCTGGTTCGTC 61
DB 96 TGGCGGCGATGACATCATCTTACCAAGCCAAAGGCTTAATGCCGCTTAGGCTCCCTC 155
QY 62 TCGTCCCTCTCTGCTGCTGGGAGAGCATCTTCAAGCTGCTGTATAGAAATTCCTT 121
DB 156 TCCCGGCTGCTGCTGCTGGGAGAGCATCTTCAAGCTGCTGTATAGAAATTCCTT 215
QY 122 GTCTTCATATTCCTCTACTATTCATCCGTGGAATCTTACAGAAATGTTCTCTGAGTAT 181
DB 216 ATCTTCTGCTCTGCTACTATCATCCGCTTATTTATAGGCTGCGCTTACCGAAGAA 275
QY 182 CAGCAGCTGTGTGTGAGAAAGTGGCTGTATCTGCAAGCTTACATTCCTCTCT 241
DB 276 CAACAGCTGATGTTGAGAAACTGACTGTATGCAAGCTTACATTCCTCTCTCTCC 335
QY 242 ATATCTCTGTTCTGCTGGTTTCTATGTATCATTTGTTGAGGCGCTGGTGGAGCCAGTAC 301
DB 336 ATTTCTCTTCTGCTGCTGGCTTCTATGCTGAGCTGCTGAGCCCGCTGGTGGAGCCAGTAC 395
QY 302 GAGAACTTGCCTGCGCCGAGCCGCTCATGATCCAGTGTCTAGCTTCTGTAGAGGCAAG 361
DB 396 GAGAACTTGCCTGCGCCGAGCCGCTCATGAGCTGCTGAGCTGCTGAGAGGCAAG 455
QY 362 GATGAGAAAGCCGTTTCTGCTGCGGAGCAGCTCATCCGCTTACGCTTCTGGGCAATG 421
DB 456 GACAGCAAGGCGCGCTGCTGCGGAGCAGCTCATCCGCTTACGCTTCTGGGCAATG 515
QY 422 CTCATCTCTGCGAGCATGAGACCTCGGCTTACAAAGGCTTCCCACTTTCACACCTG 481
DB 516 CTCATCTCTGCGAGCATGAGACCTCGGCTTACAAAGGCTTCCCACTTTCACACCTG 575
QY 482 GTGTAGCAGGTTTATGATCCATGGGAGATGAGAGTTGACAGAGTTGAGGCTTACCA 541
DB 576 GTGCAAGAGGCTTATGATCCGAGCAACAGAGTTGAGAGAACTGAGCTTACCA 635
QY 542 CACAACACATTTCTGGGTCCTTGGGTGTGTTTGCCTTGTCAATGAGGCTTATCTT 601
DB 636 CACAACATGTTCTGGGTGCTGCTGGGTGTGTTTGCCTTGTCAATGAGGCTTATCTT 695
QY 602 GGAGGTGGAATCCGAGGACACGCTGCTTCCAGAGCCGATGAGTGTATCTTGG 661
DB 696 GGAGGTGGAATCCGAGGACACGCTGCTTCCAGAGCCGATGAGTGTATGAAACCTTGG 755
QY 662 CGTACTCAGTGTGAGCAGCTGTATGCTTACAGCTGATGATGATCCATGGTGTACACA 721
DB 756 CGTACTCAGTGTGAGCAGCTGTATGCTTACAGCTGATGATGATCCATGGTGTATGACA 815
QY 722 CAGGTGTGACAGTGTGAGTATGAGCTTCTTCTGATCCTTGATGAGGAGCAGTTT 781
DB 816 CAGGTGTGACAGTGTGAGTATGAGCTTCTTCTGATGATGATGAGGAGCAGTTT 875
QY 782 CTGAACCCCAACCAAGGATACCCAGGCGATGAGTGTGATGCTGCTGCTTCA 841
DB 876 CTGAACCCCAACCAAGGATACCCAGGCGATGAGTGTGATGCTGCTGCTTCA 935
QY 842 ATCTGCAATTTCTTATCTACATGGGCTGCTGGA----- 876
DB 936 TTCTGCAAGTTCTTCTTATGTTGGCTGGCTGAAGGAGGAGGCTCTCCAGGAGCCTGCTG 995
QY 877 ----- 876
DB 996 GGCTGAGGCAATGGCCAGAGGGGTCATGGCCAGAGCTGCTGAGACGAGAGTACAGTGT 1055

```

```

QY 877 ----- 876
DB 1056 CAGGAAGAGATGTCTACCGGTTAGAAAGCAGCCAGCGGTGTGGCCGACCTGTATC 1115
QY 877 ----- 876
DB 1116 CCAGTACTCCGAGGCTGAGGAGAGATATGCTTGAACCCGAGAGCGAGAGTGTGG 1175
QY 879 TGGCAGAACAGCTATCAACCCCTTGGGAGAGAGATGATGATTTTGAAGTAACTGA 938
DB 1176 TGGCAGAGCAGCTATCAACCCCTTGGAGAGATGATGATGATTTTGAAGCAGTGA 1235
QY 939 TCATTGACAGAACTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 998
DB 1236 TTGTGACAGAAATTTGAGAGTGTCCCTGTTGGCTGTGATGATGATGATGATGATGAT 1295
QY 999 CTCGATGAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1058
DB 1296 CTCGATGAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1355
QY 1059 CTTCGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1118
DB 1356 CTTCGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1415
QY 1119 AAGAACTTGTAGCTTGTCTCAAAAGAGAGGCTGACACGGAATPAGAAAGAGTGGCT 1178
DB 1416 AAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1463
QY 1179 ATGACAGACCATAGGCTGCTTCTTGAAGCTGCAACCCAAATACTATCTTCTTGA 1238
DB 1464 AGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1523
QY 1239 AAGCTTAAAGCAAACTATGTTGTTCTAGAA-----CCCTCTCTGAGGCGAGT 1292
DB 1524 CAACCTCAAGGACCAAACTATGTTGTTCTAGAA-----CCCTCTCTGAGGCGAGT 1583
QY 1293 GTAAGGATGCCAACACAGAAACCAAGAAAGAT-----GTTGGA 1331
DB 1584 CCAAAACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCAAGGACCA 1643
QY 1332 AATTTAAGGCTGTGACTTCTTGAATGTGTTCCAGGTTTAAAGAGAGGCTTCCATT 1391
DB 1644 AGCTTAAAGGCTGTGAGGCTTCAAGCTGCCCCCAGCTATGAGGCGGAGCTTCTACA 1703
QY 1392 GTGCCCCAGAGGACCCAGGAGCCACCTTAC 1422
DB 1704 GTGCCCCAGAGGACCCAGGAGCCACCTTCC 1734

```

RESULT 6
 AA221228
 ID AA221228 standard; cDNA; 2429 BP.
 XX
 AC AA221228;
 XX
 DT 22-NOV-1999 (first entry)
 XX
 DE Human CGICE long form cDNA sequence.
 XX
 KW CGICE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
 XX age-related macular dystrophy; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 105..1412
 FT /*tag= a
 FT /product= "CGICE short form protein"
 XX
 PN W09943695-A1.
 XX
 PD 02-SEP-1999.
 XX

RESULT 7
AB061177
ID AB061177 standard; cDNA; 2112 BP.
XX
AC AB061177;
XX
DT 26-FEB-2003 (first entry)
XX
DE FLJ20132 file clone COL6441 encoding sequence.
XX
XX Neuroprotective; immunomodulator; cancer; chromosome 19; cytosolic;
KM anti-inflammation; gene therapy; nutritional supplement; wound; burn;
KM ulcer; Alzheimer's disease; Huntington's disease;
KM amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
KM vulnerability; gene; ss.
XX
OS Homo sapiens.
XX
XX MO200231111-A2.
XX
PN 18-APR-2002.
PD
XX 11-OCT-2001; 2001WO-US027760.
XX
PR 12-OCT-2000; 2000US-00687527.
XX
PA (HYSEQ-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-426278/45.
XX
PS N-PsDB; ABP43933.
XX
XX New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation.
XX
XX Claim 1; SEQ ID # 390; 357bp + Sequence Listing; English.
XX
XX The invention relates to 446 newly isolated polynucleotide sequences. The
CC activity of polynucleotides of the invention may be described as,
CC vulnerability, neuroprotective, immunomodulator, cytosolic and anti-
CC inflammatory. Compositions comprising nucleic acids of the invention are
CC useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records AB060788-
CC AB061133 represent polynucleotides of the invention. NOTE: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2112 BP; 382 A; 681 C; 607 G; 442 T; 0 U; 0 Other;

Query Match 27.1%; Score 519.4; DB 6; Length 2112;
Best Local Similarity 66.9%; Pred. No. 3,1e-121;
Matches 755; Conservative 0; Mismatches 371; Indels 3; Gaps 1;

QY 9 CCATGACTATCACTACCAAAAGTACCAATGCCGCTGGTTCGTTCTGCTCC 68
DB 190 CGATGACCGCTACCTACACAGCCGAGTGGGAAAGCCGCTTGGTGGCTTCCAGC 249

QY 69 TCTCTGCTGCTGAGGAGGAGCATCTACAGCTGCTATGAGAAATTCCTTGCTTCA 128
DB 250 TGTCTGCTACTGTGGCTGGAGCATCTACAACTCTCTGGCGAGAGCTGCTGCTTCC 309

QY 129 TATTCCTCTACTATTCATCCGTGAGACTCTACAGAAATGTTCTCTGAGTATGACGAC 188
DB 310 TTGGGTTCTATACGAGCGGTGAGTGTGCTTACCGCTTGTGCTGACCGAAGGAGGAGAC 369

QY 189 TGTGTTGAGAAAGCTGGCTCTGTACTGCGACAGCTACATTCAAGCTCATCCCTATTCCT 248
DB 370 GCTACTCTCGAAGAGCTGTGATTTATTTGTGACAGATATGCAAGCTCATCCCTGTCTCT 429

QY 249 TCGTTCGGGTTTCTATGTTACATTTGGTGTAGCCCGCTGGTGAAGCCATGACGAATCT 308
DB 430 TCGTGTGGGCTTTTATGTGACGCTGGTGAACCCCTGGTGAAGCCATGACCTATGCA 489

QY 309 TGGCGTGGCCGACCGGCTCATGATCCAGGTGTAGCTTGTGGAAGGCAAGATGAGG 368
DB 490 TGGCGTGGCCGACCGGCTCATGATGCGTGTGGGAGCACTGACGAGACGACGAC 549

QY 369 AAGCGCTTGTGCTGCGGCGACGCTCATGCGCTACGCTAGCGGCAAGTGTCTATCC 428
DB 550 GCGGCGGCTCTACCGGCGACACTCATGCGCTACGACAGGCTCTCGGCGGCTGCTCATCC 609

QY 429 TGGCGACATCAGCACTCGGTCTACAGCGCTTTCACCTCTTACCACTGCTGCTAG 488
DB 610 TGGCGTCCGTACGACCGCGGTGTTCAAGCGCTTCCCAACATAGACCACTGCTGAGG 669

QY 489 CAGGTTTATGACCCATGGGGAATATAGCATGTTCAGAAATTTGGGCTTACCAACAA 548
DB 670 CTGGGTTTATGACCCGAGAGGCAAGAAATTTGAAACCTGAACCTCATCTACACAA 729

QY 549 CATTCTGGGTCCTCGGCTGTGTTGCAACTTGTCAATGAAGGCTTCTGAGAGTC 608
DB 730 AGTACTGGGTGCTCTGCTGTGTTCTCACTGGGCGACAGGCGGACGAGGCGC 789

QY 609 GAATCCGGGACACCGCTGCTCCAGAGCTGATGAAATGAGGTGTATCTTGTGCTATC 668
DB 790 GCATCCGGAACAAGGCGCTTAAAGCTGCTCGAGAGCTAAATGTTTTTGGGGCA 849

QY 669 AGTGTGACAGCTGTATGCTTACGATCGATGATATCCATTGGTGTACACACAGGTGG 728
DB 850 AATGTGAATGCTCTTTCATCATATACGTGATTAACCTCTGCTGTACACGAGGTGG 909

QY 729 TGAAGGAGGAGTATACAGCTTTTCTTGGACCTTGATGAGGAGGAGTTTGTGAAC 788
DB 910 TGACCATGCGACTGTACAGCTTCTCTGCTTCTCTCATTTGTGCGCAATTCCTGAGC 969

QY 789 CAACAAGAGTACCCAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 848
DB 970 CGGCTCAGGGTTTCAAAAGACCAAGACCTGATGATGATGATGATGATGATGATGATGAT 1029

QY 849 AATTCTTATTTACATGGGCTGGTGAAGGTGGCAGAACAGCTATCAACCCCTTGGAGG 908
DB 1030 AGTCTTCTTCTACGCGGCTGGCTCAAGGTAGCTGAGCAAGCTCATCAACCCCTTGGAGG 1089

QY 909 AGGACGATGATGATTTTGAAGCTACATGATCATTTGACAGAAAGCTGAGGTCCTCTGT 968
DB 1090 AGGACGATGATGATTTTGAAGCTACATGATCATTTGATGATGATGATGATGATGATGATGAT 1149

QY 969 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1028
DB 1150 TGGCAGTGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1209

QY 1029 AGGACGCTCAGCGGCTCCTACAGAGCTGCTTCTG---CAGGTCTGCGCGGATTCCT 1085
DB 1210 CAGCGAGGCTCGGCGGCTCCTACAGAGCTGCTTCTG---CAGGTCTGCGCGGATTCCT 1269

QY 1086 TCATGGGCTCAGCTTCAATGACATGAGCTTAAAGAAAGAAAGCTTGAAGCT 1134
DB 1270 TCAGGAGCTCAGCTTGAATGACATGAGCTGCGCAAGAAAGAAAGCTTGAAGCT 1318

RESULT 8
ABZ80973
ID ABZ80973 standard; DNA; 2137 BP.
XX
AC ABZ80973;
XX
DT 15-OCT-2003 (first entry)

XX	DE	Human	bestrophin 2 (vitelliform macular dystrophin 2-like protein 1) gene.
XX	KW	Bestrophin; vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;	
XX	KW	antiobesity; antidiabetic; immunomodulator; hypotensive; cardiac;	
XX	KW	antiinflammatory; antidiabetic; antidiabetic; antidiabetic;	
XX	KW	energy homeostasis; metabolism; triglyceride; body-weight regulation;	
XX	KW	coronary heart disease; cachexia; diabetes mellitus; hypertension; gallstones;	
XX	KW	sleep apnea; chromosome 19.	
XX	OS	Homo sapiens.	
XX	FS	Key	
XX	FT	Location/Qualifiers	
XX	FT	CDS	
XX	FT	212..1741	
XX	FT	/tag=a	
XX	FT	/product="VMD2-like protein 1"	
XX	PN	WO2003030922-A2.	
XX	PD	17-APR-2003.	
XX	PF	09-OCT-2002; 2002WO-EP011321.	
XX	PR	09-OCT-2001; 2001EP-00124059.	
XX	PA	(DEVE-) DEVELOPMENTAL BIOLOGISCHE FORSCH.	
XX	PI	Stevener A. Broenner G, Filtsch R, Eulenberg K, Ciossek T;	
XX	DR	WPI, 2003-393411/37.	
XX	DR	P-PSDB; ABR58056.	
XX	PT	New pharmaceutical compositions comprising a Bestrophin gene, polypeptide	
XX	PT	or nucleic acid, for treating, alleviating and/or preventing metabolic	
XX	PT	diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or	
XX	PT	gallstones.	
XX	PS	Claim 2; Fig 4c; 85dp; English.	
XX	XX	This sequence represents the coding region for a member of the human	
XX	CC	bestrophin gene family designated vitelliform macular dystrophin 2 (VMD2)	
XX	CC	-like protein 1. The dystrophin gene family are involved in energy	
XX	CC	homeostasis and metabolism of triglycerides. The sequence can be used for	
XX	CC	the manufacture of an agent for detecting and/or verifying, for treating,	
XX	CC	alleviating and/or preventing disorders including metabolic diseases such	
XX	CC	as obesity and other body-weight regulation and related disorders such	
XX	CC	as eating disorder, cachexia, diabetes mellitus, hypertension, coronary	
XX	CC	heart disease, hypercholesterolemia, osteoarthritis, gallstones, cancers	
XX	CC	of the reproductive organs, and sleep apnea. The gene is found on human	
XX	XX	chromosome 19	
XX	XX	Sequence 2137 BP, 399 A; 671 C; 622 G; 445 T; 0 U; 0 Other;	
XX	XX	Query Match	
XX	XX	Best Local Similarity 27.1%; Score 519.4; DB 8; Length 2137;	
XX	XX	Matches 755; Conservative 0; Mismatches 371; Indels 3; Gaps 1	
XX	QY	9 CCAAGACTATCACTACCAACAAAGATGACCAATGCCCGCTGGTTCCTGCTCC 68	
XX	DB	210 CGAATGACCGTACCACTACCAACAGCCAGAGGCGCAAGCCCGCTTCGATCTCCACAGC 269	
XX	QY	69 TCCCTCTGTGCGGAGGAGGAGCATCTCAACAGCTGTGATGAGAAATCCCTTGTCTCA 128	
XX	DB	270 TGCCTCTACTGTGGGTGGAGCATCTCAACACTCTGTGGGAGAGCTGCTCTGCTTC 329	
XX	QY	129 TATTCCTTACTATTCATCCGTGACTCTACAGAAATGTTCTTCGATGATCAAGC 188	
XX	DB	330 TTGGGATCTACATGCGCTGATGCTGCTACCCGCTTGTGCTACCGAAGGCGAAGC 389	
XX	QY	189 TGTGTTTGAAGAGTGGCTGTATCTGACAGAGTATCAATTCAGTATCTCTATATCTT 248	
XX	DB	390 GCTACTCTGAGAAAGTTGTGATTTATTTATGTGACAGATTAAGCCAGCTCAATCCCTGTCTCT 449	

[illegible]

XX	antihypertensive; antidiabetic; immunomodulatory; hypotensive; cardiant;
KW	antihypertensive; osteoporathic; antiinflammatory; cytostatic; obesity;
KM	energy homeostasis; metabolism; triglyceride; body-weight regulation;
KW	eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
KM	coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
XX	sleep apnea; chromosome 19.
XX	
OS	Homo sapiens.
XX	
XX	Koy Location/Qualifiers
PH	1..1530
FT	CDs
FT	/*tag= a
FT	/product= "VMD2-like protein 1 homologue"
XX	
XX	MO2003030922-A2.
XX	
PD	17-APR-2003.
XX	
XX	09-OCT-2002; 2002MO-EP011321.
XX	
PR	09-OCT-2001; 2001EP-00124059.
XX	
PA	(DEVE-) DEVELOPEN ENTWICKLUNGSBIOLOGISCHE FORSCH.
XX	
PI	Steuernagel A, Broenner G, Fritsch R, Eulenber K, Ciosek T;
DR	WPI; 2003-393411/37.
DR	P-PSDB; ABR58059.
XX	
PT	New pharmaceutical compositions comprising a Beetrophin gene, polypeptide
PT	or nucleic acid, for treating, alleviating and/or preventing metabolic
PT	diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
XX	galactones.
XX	
PS	Disclosure; Fig 41; 85pp; English.
XX	
CC	This sequence represents the coding region for an alternative member of
CC	the human beetrophin gene family designated vitelliform macular
CC	dystrophin 2 (VMD2)-like protein 1. The dystrophin gene family are
CC	involved in energy homeostasis and metabolism of triglycerides. The
CC	sequence can be used for the manufacture of an agent for detecting and/or
CC	verifying, for treating, alleviating and/or preventing disorders
CC	including metabolic diseases such as obesity and other body-weight
CC	regulation and related disorders such as eating disorder, cachexia,
CC	diabetes mellitus, hypertension, coronary heart disease,
CC	hypercholesterolemia, osteoarthritis, galactones, cancers of the
CC	reproductive organs, and sleep apnea. The gene is found on human
CC	chromosome 19
XX	
XX	Sequence 1530 BP; 257 A; 493 C; 462 G; 318 T; 0 U; 0 Other;
XX	
XX	Query Match 27.1%; Score 519; DB 8; Length 1530;
XX	Best Local Similarity 66.9%; Pred No. 3.4e-121;
XX	Matches 754; Conservative 0; Mismatches 370; Indels 3; Gaps 1
XX	
QY	11 ATGACTATACCTACACAAACAAAGTAGCCATGCCCGCTGGTTCGTTCTGTCCTC 70
Db	1 ATGACCGTCACCTACACAGCCCGAATGGGAAACGCCCGCTGGTGGCTTCTCCACGCTG 60
QY	71 CTCTGTGCTGGCCGAGGACGATCTACAAAGCTGCTGTATGAGAAATTCCTTGTTCATA 130
Db	61 CTGCTACTGTGGCCGGAGGACATCTACAAACTCTGTGGCGAGACTGCTGTGCTTCTT 120
QY	131 TTCTCTACTATTCATCCATCCGTGAGACTCTACAGAAATGTTCTCTCGAGTACGAGAGCTG 190
Db	121 GGGTTTTCATAGGCGCTGAGTCTGCTTACCGCTTTGTGCTATCCGAAAGGAGAAAGCC 180
QY	191 TTGTTTGAAGAGCTGGCTCTGTACTGCGACAGCTACATTACAGCTCATCCCTTATCTTC 250
Db	181 TACTTCGAGAAAGCTGTGATTTATTTGTGACAGTATGCGAGCTCATCCTGTCTCTTC 240
QY	251 GTTCGGGTTCTATGTTACATTTGTTGGTGGAGCCGCTGTGAGCCAGTACGAAACTTG 310

Db	241	GTGCTTGAGCTTTTATGAGACGCTGTGAGTAAACCGCTGTGTGAGACAGTAACTTATGCATG	300
Qy	311	CCGTGGCCCGAACCCTCATGATCCAGGTGCTTACGTTCTGTGGAGGGCAGAGATGAGGAA	370
Db	301	CCGCTGCCCGACGCGCTCATGTGTGTGGTGGCGGGCACCCGTGCACCGACCGACGCCG	360
Qy	371	GGCGGTTTGTGGGGGGCAGCGCTATCGGCTGACGCCATCCGTGGGGCAAGTGTCTATCCGT	430
Db	361	GGCGGCTCTTACCGGGCGACACTATGTGGCTACGCGAGGGGCTCTGGCCGTCTATCCGTG	420
Qy	431	CGCAGCATACGACACTGTGCTTACAGCGCTTTCCACTTCTTCAACACCTGTGTACGA	490
Db	421	CGCTCCGTGACGACCGGGGTCTTCAAGCGCTTCCCGACATAGACACAGTGTGAGGCT	480
Qy	491	GGTTTATAGACCCATGGGGAAATATAGCAGTTGAGAGATTGGGCTTACCAACACACACA	550
Db	481	GGGTTTATAGACCCGCGAGGAGCGCAGAAAGTTTGAAAACCTGAATCTACTTACACAAG	540
Qy	551	TTCTGGGTGCCCTGGGTGTGGTTTGCCACTTGTTCATGAAAGGCTATCTTTGAGTGTGA	610
Db	541	TACTGGGTGCCCTGGGTGTGGTTTGCCACTTGTTCACACTTGGCGCACAGGGCCGACGGAGGGCGC	600
Qy	611	ATCCGGGACACCGCTGCTGTCCAGAGCGCTGATGTAATGAGGTGTATCTTTGCTGTA	670
Db	601	ATCCGGGACACACCGCGCTTAAAGCTGTGCTGTCCAGAGAGCTGTAATGTTTTTGGGGGCAA	660
Qy	671	TGTGGACAGCTGTATGCTTACGACTGTGATATCCCATTTGTGTACACAGGTGTG	730
Db	661	TGTGGATGTCTTTCATCTATGACTGTGATAGCTAACCCCTCGTGTACACAGAGTGTG	720
Qy	731	ACAATGGCAGTATACAGCTTTTCTTGCATGTCTTATCGGGAAGCAGTTTCTGAACCCA	790
Db	721	ACCAATGGCAGTATACAGCTTTCTTGCATGTCTTGCATGTGTGCGCAGTTTCTGGAACCG	780
Qy	791	AACAAGAGCTAACCCAGGCGCATGAGATGGATCTGTGTGCTGTCTTCACAATCTTGCA	850
Db	781	GCTAGAGGTATCAAAAGACACAGACTTGAACCTGTGTGTGCCCATTTCAACCTTTTGAG	840
Qy	851	TTCTTATTTCTACATGAGGCTGTGTGAGGTGGCAGAAAGCTCATCAACCCCTTGGGGAG	910
Db	841	TTCTTCTTTCTTACGCGGGCTGGCTCAAGGTAGCTGACGCTCATCAACCCCTTGGGAAG	900
Qy	911	GACGATGATGATTTTGAAGCTAACTGATCTTATGACAGAAACCTGACAGGTGTCCCTGTTG	970
Db	901	GACGATGATGATTTTGAAGCAACTTTCATGATGATGAACCTTCAAGGTGTCCATGCTG	960
Qy	971	TTCGTGATGAGGATGACACAGAAATTGCTCCCATGGAACGTGACATGTATGTGAACGAG	1030
Db	961	GCACTGTGACAGAGATGTATGATGACTGTGTGTGGAGAAAGACTGTATCTGGAGATCA	1020
Qy	1031	GCAGCGGCTCAGAGCGGCTCTACACAGCTGCTTCTG---CCAAGTCTCGCGGCAATTCCTTC	1087
Db	1021	GCCGAGGCTCGCGGCTCATACAGCGGCTACTGTCTTCCAGCTGCGGCAAGCTTTCCTTC	1080
Qy	1088	ATGGGCTTCACTTCAACATCAGCTTAAAGAAAGAACTTGAAGCT	1134
Db	1081	CAGGGCTCAACCTTGAACATCACGCTGCGCAAAAGAAACATGCAGTT	1127
RESULT 10			
ADMO2614			
XX	ID	ADMO2614	standard; cDNA; 2500 BP.
XX	AC	ADMO2614;	
XX	DT	20-MAY-2004	(first entry)
XX	DE	Human cDNA of the invention SEQ ID NO:1299.	
XX	XX	ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.	
XX	OS	Homo sapiens.	
XX	XX		

PN EP1347046-A1.
XX
PD 24-SEP-2003.
XX
PF 12-APR-2002; 2002EP-00008400.
XX
PR 22-MAR-2002; 2002JP-00137785.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
DR WPI, 2003-723558/69.
XX P-Psdb; ADM05057.
XX
PT New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.
PS
PS Claim 1; SEQ ID NO 1299; 305pp; English.
XX
XX The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC cDNA sequence of the invention.
XX
SQ Sequence 2500 BP; 689 A; 632 C; 563 G; 616 T; 0 U; 0 Other;
Query Match 23.9%; Score 458.2; DB 11; Length 2500;
Best Local Similarity 63.5%; Pred. No. 1.2e-105;
Matches 700; Conservative 0; Mismatches 403; Indels 0; Gaps 0;
QY 11 ATGACTATACCTTACACAAAGTAGCCCAATGCCGCCCTCGGTGCTTCTGCTCCCTC 70
Db 229 ATGACTGTCATCTTACCTCAATAGTAGCAATAGCACTTTTGTGATTCATAGTTTA 288
QY 71 CTCCTGTCGTCGCGAGGAGCATCTACAGAGCTGCTGATGAGAAATTCCTTCTTCATA 130
Db 289 CTCCTCAAGTGAAGAGGAGCATCTACAACTACTGACAGGAATTTATGTTTCT 348
QY 131 TTCCTTAATAATTCATCCGTCGACTCTACAGAAATGTTCTCTGAGTATCAGCAGCTG 190
Db 349 GTTCTTTATACAGCAATTAAGTTTGTGTGACAGATTGTTACTTACAGAGTCCAAAACGT 408
QY 191 TTGTTTGAAGAGTGGCTCTGACTGAGACGACATCAATCGTCATCCCTAATCCCTTC 250
Db 409 TACTTTGAAAATTAATATCAATTACTGTGACAGATATGCTGAACAAAATTCAGTAACTTT 468
QY 251 GTTCTGAGTTTCTATGTTTACATTGGTGGTGAAGCGCTGTCGTCGACAGCAACTTG 310
Db 469 GTGCTTGGGTTTATGTTTACTCTGTGATGAAACCGATGGTGAACCAAGTTTGTGAATTG 528
QY 311 CCGTGCCCGGACCGCCTCATGATCCAGGTGTCTAGCTTCTGTGAGGGCCAAAGATAGGAA 370
Db 529 CCGTGCCCGGACAGAGGCTTAATGTTCTCATCTCTAGCAGTGTTCACGGAAGAGAGAC 588
QY 371 GGGCGTTTGTGGGGGCGGCGCATCGCTACCGCCATCCCTGGGCCAAGTGTCTATCTG 430
Db 589 GGGCGCTTGTCTTAAGAGAGCGCTGATGGCTACCTCATCTCACTCCCTGCTATCTTT 648
QY 431 CGCAGCATCAAGCCTCGGCTTCAAGAGCGCTTTCCTTCCACCTGCTGTCAGCA 490
Db 649 CGCTCGGTGAGCATGCTGTGTGACAAAAGATTTCCCATGATGACACAGTGTGTAAGCA 708
QY 491 GGTTTTATGACCCATGGGGAACATAGCAAGTTGAGAAAGTTGGGCTTACCAACAACACA 550

Db 709 GGTTTATGACACAGATGAAGAAATTAATTCACACACTCAAGTCTCTCATCTGAAA 768
QY 551 TTCTGGGTGCTCCGTGGGTGTGTTGCCAATCTGTCAATGAAGCTTATCTTGAGTCA 610
Db 769 TATGGGTTCATTCATCTGTGTGAAATCTTGCAATCAAGCCGGAATGAAGTGA 828
QY 611 ATCCGGAGACCCGTCCTGTCAGAGCTGATGAATGAGTGTATCTTGCCGACTCAG 670
Db 829 ATCAGACAGGTGTTGATCTGCATATCTAATGACTGAATGAATCATACCGCTTGG 888
QY 671 TGTGACAGCTGATGCTTACGACACTGATGAATGCCATTTGTGTATACACAGGTGTG 730
Db 889 TGCAGCTCTTATTCGCTGTATGACTGTGGATTCGCTGCTTACACCCAGTGTGC 948
QY 731 ACACTGACATATACAGCTTTTCTTCTGATGCTGATTCGGAGAGCAGTTTCAACCA 790
Db 949 ACTTGTCTGTATATACCTTCTTCTTGTGCGTGTGATGAGCGCCAGTTTGGATCCC 1008
QY 791 AACAAAGACTAACCCAGCCATGAGATGATCTGTGTGCTGCTGCTTCAACATCTGCA 850
Db 1009 ACCAAAGCTAACCGAGGCAATGACTTGTGATCTTCAATTCCTTCACTTCTTCA 1068
QY 851 TTCTTATCTTACATGCGCTGCTGAAAGTGGCAGAAACAGCTCATCAACCCCTTCG 910
Db 1069 TTCTTCTTATGAGAGATGCTTAAAGTACAGAGAGCTTATCAACCTTTTGGAGAA 1128
QY 911 GAGCATATATTTTGAAGATTAATCTGATATTTGACAGAAACCTGCAAGTGTCCCTG 970
Db 1129 GATGATGATATTTTGAATTAATCTGATGATTTGACAGAAATTTGAGAGGCTCT 1188
QY 971 TCGGTGATGAGATGACACAGAACTTGCCTCCCATGGAAGTATGATCTGGAACGAG 1030
Db 1189 GCTGTGACAAATGACATGAGCTTACCCAAATGAAGAAKATTTATGAGAGAT 1248
QY 1031 GCAGCGCTGAGCGGCTTACACAGCTGCTTGCAGAGTCTGCGCGCATTCCTTCATG 1090
Db 1249 TGTGCTGCTGCGCCACCATATCATTTGAGCAGTGTGACTGATCATATCCCTATTT 1308
QY 1091 GGCTCCACCTTCAACATCAGCCT 1113
Db 1309 GGGTCAACAGTCCAGATGGGGCT 1331
RESULT 11
AB280974
ID AB280974 standard; DNA; 2028 BP.
XX
AC AB280974;
XX
DT 15-OCT-2003 (first entry)
XX
DE Human beetrophin (vitelliform macular dystrophin 2-like protein 3) gene.
XX
XX Bestrophin; vitelliform macular dystrophin 2; VMD2; ds; gene; vaccine;
XX antiobesity; antidiabetic; immunomodulator; hypotensive; cardiant;
XX antilipemic; osteopathic; antiinflammatory; cyostatic; obesity;
XX energy homeostasis; metabolism; triglyceride; body weight regulation;
XX eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
XX coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
XX sleep apnea; chromosome 12.
OS
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..2028
FT CDS /tag= a
FT /product= "VMD2-like protein 3"
XX
XX MO2003030922-A2.
XX
PD 17-APR-2003.
XX

Query Match	22.1%	Score 423	DB 8	Length 2038
Best Local Similarity	62.2%	Pred. No. 9-97		
Matches 697	Conservative 0	Mismatches 400	Indels 24	Gaps 1
09-OCT-2002; 2002MO-EP011321.				
09-OCT-2001; 2001EP-00124059.				
(DEVE-) DEVELOGEN ENTWICKLUNGSBIOLOGISCHE FORSCH.				
Steuernagel A, Broenner G, Fritsch R, Eulenberg K, Ciosek T;				
WPI; 2003-393411/37.				
P-PsDB; ABR58057.				
New pharmaceutical compositions comprising a Beetrophin gene, polypeptide or nucleic acid, for treating, alleviating and/or preventing metabolic diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or gallstone.				
Claim 2; Fig 4e; 85pb; English.				
This sequence represents the coding region for a member of the human beetrophin gene family designated vitelliform macular dystrophin 2 (VMD2)-like protein 3. The dystrophin gene family are involved in energy homeostasis and metabolism of trisylsterides. The sequence can be used for the manufacture of an agent for detecting and/or verifying, for treating, alleviating and/or preventing disorders including metabolic diseases such as obesity and other body-weight regulation and related disorders such as eating disorder, cachexia, diabetes mellitus, hypertension, coronary heart disease, hypercholesterolemia, osteoarthritis, gallstones, cancers of the reproductive organs, and sleep apnea. The gene is found on human chromosome 12				
Sequence 2028 BP; 543 A; 511 C; 463 G; 511 T; 0 U; 0 Other;				
11 ATGATCTACCTTACACAAACAAAGTACGATGCCGCTGGTTCGTCTCTCCCTC 70				
1 ATGATCTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 60				
71 CTCCTGCTGCTGGCGAGGAGCATCTACAAAGCTGTATGAGAAATTCCTTCTCATTA 130				
61 CTCCTCAAGTGGAGGAGGAGCATCTACAAAGCTGTATGAGAAATTCCTTCTCATTA 120				
131 TTCCTCTACTATTCATCCGCTGAGACTTACAGAAATGTTCTCGAGATACAGCAGCTG 190				
121 GTTCTTTATACAGCAATTAAGTTGGATATACAGATTTGTTACTTACAGAGTCCAAAAGCT 180				
191 TTGTTTGAAGAACTGGCTCTGTACTGCGACAGCTTACATTCAGCTCATCTTATCTTC 250				
181 TACTTTGAAAATTTCAATTTACTGCTGACAGATATGCTGAAACAAATTCAGTAACTTT 240				
251 GTTCTGGGTTTCTATGTTACATTTGATGGAGACCGCTGGTGGAGCAGTACAGAACTTG 310				
241 GTGCTGGGTTTATGTTACTTGGATGTAACCAATGGTGAACCAAGTTTGTGAATTTG 300				
311 CCGTGGCCGACCGCTCATGATTCAGGCTGTACTTCTGTGAGAGGCAAGATGAGNA 370				
301 CCGTGGCCGACCGCTCATGATTCAGGCTGTACTTCTGTGAGAGGCAAGATGAGNA 360				
371 GGGCGTTTGTGCGGCGACGCTTACCGCTACGCGATCTTGGGCGAAATGCTCATCTCTG 430				
361 GGGCGTTTGTGCGGCGACGCTTACCGCTTACCGCTTACCGCTTACCGCTTACCTTT 420				
431 CGCAGCATCAGACCTCGGTCTACAAAGGCTTTCCCATCTTACACCACTCGGTCTAGGA 490				
421 CGCTCGGTAGAGCATCTGTGTACAAAGATTTCCACAAATGAGACCATGTTGTAAGA 480				
491 GGTTTTATAGACCATGAGGAGCATAGAGCTTACAGAAATTTGGGCTTACCAACACACA 550				
481 GGTTTTATAGACCATGAGGAGCATAGAGCTTACCAACACACTCAAGTCTCTCATCTGAA 540				
551 TTCTGGGTGCTCGGTGTGTTTGGCACTTGTCAATGAGGCTTATCTTGGAGTCTGA 610				

Db	541	TATTGGGTTCATTCATCTGGTGTGGAAATCTTGGAACTTAAGCCGGAAATGAAGTGA	600
Oy	611	ATCCGGGACACCGTCCTGCTCCAGAGCTGATGAATGAGGTGTACTTTGGCTAC	670
Db	601	ATCAGAGACAGATGTGATCTGCATTCATTTGATGACTGAAATGAAATCAGATACCGCTTTGG	660
Oy	671	TGTGGACAGCTGTATGCTTACGACTGGAATAGTATCCATTTGCTGTACACACAGGTGGG	730
Db	661	TGCAGCCTTATTTTCGGTTATGACTGGGTGGGATTCGGCTGTTTACACCCAGGTGTG	720
Oy	731	ACAGTGCAGATATACAGCTTTTCTCTTCATGCTTGATTCGGAGGACGTTTCTGAACCA	790
Db	721	ACTCTTGCTGTATACCTTCTTCTTTCGCGCTCGATTTGACGCGCACTTTTGGATCC	780
Oy	791	AACAAGACTACCCAGGCGCATGAGATGAGATCTGTTGTGCTGTCTTCAATCTTGCA	850
Db	781	ACCAAGGCTACGACGAGGACATGACTTGGATCTTAACTTCCATCTTCAACCTCTTAA	840
Oy	851	TTCTTATTTCTACATGGGCTGCTGAG-----GTGGCAGAA	886
Db	841	TTCTTCTTCTATGACAGATGGCTTAAAGATTTATATATATGTACTAAAGTAGTACAG	900
Oy	887	CAGCTCATCAACCCCTTGGGGGAGAGCATGATGTTTGTAGACTTACATCTGATCAATTG	946
Db	901	CAGCTTATCAACCTTTTGGAGGAATATATGATTTTGAATCTTACATCTGCTGATTA	960
Oy	947	AGAAACCTGCAGGTGTCCCTGTTGTCGTGATGGGATGACACAGAACTTGCTCCATG	1006
Db	961	AGAAATTTGACAGGTCTCTTTTATGCTGTGACAGAAATGACATGAGCTTACCAAGATG	1021
Oy	1007	GAACTGACATATCTGAGATGAGGACAGCGCCTCAGCGCCCTTACACAGCTGTCTTGC	1066
Db	1021	AAGAAGGACATTTTCTGGGAGATTTCTGCTCTCGCCACCATATCACATTGGCAGCTGT	1086
Oy	1067	AGGTCTGCGCGGCACTTCCTCATGGGCTCCACCTTCAACAT	1107
Db	1081	GACTACTGACATACCTCTCATTTCTGGGGTCAACAGTCA	1121
RESULT 12			
ACCS9900			
ID	ACCS9900 standard; cDNA, 1948 BP.		
AC	ACCS9900;		
XX			
DT	07-JUL-2003 (first entry)		
XX			
DE	Human REMAP-13 encoding cDNA SEQ ID NO:49.		
XX			
KW	Human; receptor and membrane-associated protein; REMAP; cytosolic;		
KW	antiartherosclerotic; anticoagulant; neurotropic; neuroprotective;		
KW	thymoprotective; anti-HIV; antiallergic; anti-inflammatory; cancer;		
KW	thyromimetic; gene therapy; cell proliferative disorder; atherosclerosis;		
KW	neurological disorder; epilepsy; Huntington's disease; stroke; AIDS;		
KW	immune disorder; inflammatory disorder; allergy; developmental disorder;		
KW	hypothyroidism; Cushing's syndrome; infection; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	333..1754	
FT		/tag= A	
FT		/product= "REMAP-13"	
XX			
PN	WO2003025130-A2.		
XX			
PD	27-MAR-2003.		
XX			
PF	12-SEP-2002; 2002WO-US029220.		
XX			
PR	14-SEP-2001; 2001US-0322157P.		
PR	28-SEP-2001; 2001US-0326029P.		

PR 05-OCT-2001; 2001US-0327380P.
 PR 12-OCT-2001; 2001US-0329198P.
 PR 19-OCT-2001; 2001US-0343742P.
 PR 02-NOV-2001; 2001US-0343906P.
 PR 02-NOV-2001; 2001US-0343980P.
 PR 16-NOV-2001; 2001US-0332426P.
 PR 13-MAR-2002; 2002US-0364388P.
 PR 15-MAR-2002; 2002US-0364494P.
 PR 29-MAR-2002; 2002US-0369248P.
 XX
 BA (INCYTE GENOMICS INC.
 XX
 PI Warren BA, Gietzen KJ, Lai PG, Xu Y, Tian UK, Lee S;
 PI Borowsky ML, Becha SD, Sanjanmala MM, Thangavelu K, Tang YF;
 PI Honcheil CD, Yue H, Elliott VS, Richardson TW, Azimzai Y, Chavla NK;
 PI Baughn MR, Lu DM, Nguyen DB, Kalafus DP, Saperstein SK;
 PI Rankum J, Lehn-Mason PM, Griffin JA, Duggan BW, Lee ST;
 PI Zebardjian Y, Hatalia AJA, Gururajan R, Swarnakar A, Lee EA;
 PI Marquis JP, Khare R, Emerling BW, Jhang X, Jackson AA;
 XX
 DR MPI; 2003-354596/33.
 P-PSDB; ABR43181.
 XX
 PT New human receptors and membrane-associated proteins (REMAP), useful for
 PT diagnosing, treating and preventing diseases or conditions associated
 PT with the aberrant REMAP expression e.g. cancer, AIDS, atherosclerosis, or
 PT infections.
 XX
 PS Claim 5; Page 224-225; 241pp; English.
 XX
 CC ACCS9888 to ACCS9923 encode the human receptor and membrane-associated
 CC proteins given in ABR43169 to ABR43204 and designated REMAP-1 to REMAP-36
 CC (1). (1) have cytosolic, antiarteriosclerotic, anticonvulsant,
 CC neurotropic, neuroprotective, cerebroprotective, anti-HIV, antiallergic,
 CC antiinflammatory and thymimetic activities, and can be used in gene
 CC therapy. The REMAP polypeptides and polynucleotides are useful in
 CC diagnosing, treating and preventing diseases or conditions associated
 CC with the decreased expression or overexpression of REMAP, such as cell
 CC proliferative (e.g. cancer, atherosclerosis), neurological (e.g.
 CC epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS,
 CC allergies) and developmental (e.g. hypohydroidism, Cushing's syndrome)
 CC disorders, or infections. They are also useful in assessing the effects
 CC of exogenous compounds on the expression of nucleic acid and amino acid
 CC sequences of REMAP. The REMAPs or their fragments are useful in screening
 CC compounds for effectiveness as agonist or antagonist of the polypeptides,
 CC or in altering the expression of the target polynucleotide and compounds
 CC that specifically bind to or modulate the activity of the polypeptide
 CC
 SO Sequence 1948 BP; 371 A; 622 C; 564 G; 391 T; 0 U; 0 Other;
 Query Match 20.9%; Score 401.2; DB 9; Length 1948;
 Best Local Similarity 61.2%; Pred. No. 3.2e-91;
 Matches 711; Conservative 0; Mismatches 403; Indels 48; Gaps 2;
 QY 9 CCATGACTATACCTACCAAAAGTACCAATGCCCGCTGCTTCTCTGCC 68
 DB 331 CCATGACGGTTTCATACCTCTCAAGTGGCGAGGCCCGCTTGGAGTTCTCTGCC 390
 QY 69 TCCCTGCTGCTGGCGAGGCGCATCTACAGCTGCTGTATGAGGAATTCCTTCTCA 128
 DB 391 TGCCTTCCTCGCTGAGGGGAAAGCATCTACAGCTCTCTACAGGAATTCCTCTTTG 450
 QY 129 TATTCCTCTACTATTCATCCGTGAACTCTACAGAAATGTTCTCTCGAGTATCAGCAGC 188
 DB 451 GGGGCTTGTAACGCTGTGCTTATGATCACTACCGGCTGCTCTACCCAGAGAGAGGT 510
 QY 189 TGTGTTTGAAGAGCTGCTGTATCTGCAAGCTACATTCAGCTCATTCCTTATCT 248
 DB 511 ACGGTATGCTCAGATGCGCCGTAATGCAACCGCTCAGCAGACCTATTCCTTGTCT 570
 QY 249 TCGTTCGGGTTTCTATGTTACATGTTGGTGAAGCCGTGTGAGCAGCAGTACGAACT 308
 DB 571 TTGATTGGGTTTCTATGTAAGCTCTCGTGTGAACCGGTGTGTCTCCAGTACACAAGCA 630

QY 309 TGCCGTGGCCCGACCGCTCATGATCCAGGTGTCTAGCTTGTGAGGCGAAGATGAG 368
 DB 631 TCCCGCTGCGAGACCACTGATGTGCTCATCTCGGCTAGGTGACAGCGCTGACCAAGC 690
 QY 369 AAGCGCTTTGCTGCGGCGCAAGCTCATCCGTACAGCCATCTGGGCGAAAGTCTCATTC 428
 DB 691 GGGGCGCGCTGCTGCGCGCAAGCTCATCCGTACCGGAACCTTGGCTGCTGCTGTC 750
 QY 429 TGCCGACATCAGACCTCGGTCTACAAAGGCTTTCCCACTCTTACCACTCGTGTAG 488
 DB 751 TGCGCTGCTCAGACCGCGGTCTTAAGCGCTTCCACATGACAGCTGTGTGACG 810
 QY 489 CAGGTTTATGACCATGAGGGAACATTAAGACGTTGCGAAGTTGGGCTTACACACAACA 548
 DB 811 CAGGTTTATGATCCCGAGGAAGAGGAAAAAGTTTGAAGCCTTAATTCGACTTCAACA 870
 QY 549 CATTCGTGGTCCCTGGGTGTGTTGGCCAACTTGTCAATGAAGGCTTATCTTGAAGTC 608
 DB 871 AGTACTGGGTCCCTGCGTCTGGTTCAACAAGCTGGCGCCAGGCCCGGAGGGAAGGCG 930
 QY 609 GAATCCGGGACACCGCTGCTCCAGGCGCTGATGAATGAAGTGTATCTTGTGATCTC 668
 DB 931 GAATACGTGACGATATGCTCTCTGTCTACTTTGGAAGAGCTGAACAGTACCGAGCCA 990
 QY 669 AGTGTGACAGCTGTATGCTTACGACTGATTAAGTATCCCATTTGTTGTAACAAGGTGG 728
 DB 991 AGTGCACATGATCTATTCATATGATGATGATGATGATGATGATGATGATGATGATG 1050
 QY 729 TGACAGTGGCAGTATACAGCTTTTCTTGCATGCTTGTATGATGAGGAGGCTTCTGAAC 788
 DB 1051 TGACATGAGCGGTCTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 826
 QY 789 CAACAGAGAGCTACCCAGGCGATGAGATGATGATGATGATGATGATGATGATGATG 826
 DB 1111 CAGAGCAGGAGGCTGCCAATCTGAGAGCTTGAAGCCAGGCGAAGCAGCCAG 1170
 QY 827 -----GTGCTGTCTTCAACATCTGCAATTTCTTATTTTACA 863
 DB 1171 CCCTGGAGACCCGAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 1230
 QY 864 TGGGCTGTGAGAGGTGCGAAGCAAGCTCATCAACCTTGGGAGAGAGATGATGAT 923
 DB 1231 CTGGCTGTGCAAGGTGCTGAAAGATCATCAACCATTTGTGTGAGATGATGATGATG 1290
 QY 924 TTGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 983
 DB 1291 TTGAGACAAATCAGCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1350
 QY 984 TGACACGAAGCTTGCCTCCATGGAACGTGATGATGATGATGATGATGATGATGATGATG 1043
 DB 1351 TGTATCGAAGACCTTCCCGCGTGAAGAGCAAGTATGATGATGATGATGATGATGATG 1410
 QY 1044 CGCCCTTACAC---AGTGTCTTTCGCAAGGTCTGCGGCGCATTCCTTATGAGGCTCAACT 1100
 DB 1411 CACCCCTAACAATGTGGGCAACGCGCGAGCTGTGGGCGCTCATTCCTGGCTTCAACT 1470
 QY 1101 TCAACATCAGCTTAAGAAAGA 1122
 DB 1471 TCAACTGCGCATGAGGAGCA 1492
 RESULT 13
 AB280975
 ID AB280975 standard; DNA; 1422 BP.
 AC AB280975;
 AC
 AC
 DT 15-OCT-2003 (first entry)
 XX
 DE Human bestrophin (vitellogenin macular dystrophin 2-like protein 2) gene.
 XX
 KW Bestrophin; vitellogenin macular dystrophin 2; VMD2; ds; gene; vaccine;

KM antiobesity; antidiabetic; immunomodulator; hypotensive; cardiac;
KM antilipemic; osteopathic; antiinflammatory; cyostatic; obesity;
KM energy homeostasis; metabolism; triglyceride; body-weight regulation;
KM eating disorder; cachexia; diabetes mellitus; hypertension; gallstones;
KM coronary heart disease; hypercholesterolemia; osteoarthritis; cancer;
KM sleep apnea; chromosome 1.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT CDS 1..1422
FT /tag= a
FT /product= "VMD-2-like protein 2"
XX MO2003030922-A2.
XX 17-APR-2003.
XX 09-OCT-2002; 2002MO-EP011321.
XX 09-OCT-2001; 2001EP-00124059.
XX (DEVE-) DEVELOPENTWICKLUNGSBIOLOGISCHE FORSCH.
XX Steuernagel A, Broenner G, Fritsch R, Eulenberg K, Ciosek T;
XX WPI: 2003-393411/37.
XX P-P5DB: ABR58058.
XX
XX New pharmaceutical compositions comprising a Bestrophin gene, polypeptide
PT or nucleic acid, for treating, alleviating and/or preventing metabolic
PT diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
PT galactones.
XX
XX Claim 2; Fig 4g; 85pp; English.
XX
XX This sequence represents the coding region for a member of the human
CC bestrophin gene family designated vitelliform macular dystrophin 2 (VMD2)
CC -like protein 2. The dystrophin gene family are involved in energy
CC homeostasis and metabolism of triglycerides. The sequence can be used for
CC the manufacture of an agent for detecting and/or verifying, for treating,
CC alleviating and/or preventing disorders including metabolic diseases such
CC as obesity and other body-weight regulation and related disorders such as
CC eating disorder, cachexia, diabetes mellitus, hypertension, coronary
CC heart disease, hypercholesterolemia, osteoarthritis, gallstones, cancers
CC of the reproductive organs, and sleep apnea. The gene is found on human
CC chromosome 1
CC
XX Sequence 1422 BP; 258 A; 468 C; 413 G; 283 T; 0 U; 0 Other;
SQ
Query March 20.8%; Score 399.2; DB 8; Length 1422;
Best Local Similarity 61.1%; Pred. No. 9.1e-91;
Matches 709; Conservative 0; Mismatches 403; Indels 48; Gaps 2;
QY 11 ATGACTATACCTTCACAAACAAAGTAGCCATGCCGCTCGGTGTTCTCGTCCCTC 70
Db 1 ATGACGGTTTATATACATCTCTCAAGTGGCGGAGCCGCTTCGAGGTTTCTCGGCTG 60
QY 71 CTCCTGCTGCGGCGAGGAGCATCTACAGCTGTGATGAGAAATTCCTTGTCTTCA 130
Db 61 CTCCTCGCTGAGGAGGAGCATCTACAGCTCTCTCAAGAAATTCCTCTTGGG 120
QY 131 TTCCTTACTATTCATCCGTGAGCTCTACAGAAATGTTCTCTGAGTATCAGACGTC 190
Db 121 GCTTGTACGCTGTGCTTACATCACTACCGGCTGCTGACCCAGAGCAAGATAC 180
QY 191 TTGTTTGAAGAGCTGCTCTGACTGCGACAGATTACATTCACTCATATCTTC 250
Db 181 GTGATGCTCAGTGCGCCGCTACTGCAACGCTCAGAGACCTCATCTCTTCTT 240
QY 251 GTTGTGGTTTCTATGTTTACTATTGGTGTGAGCCGCTGTGTGAGGACAGAACTTG 310
Db 241 GTATTGGGTTTCTATGTGACTCTGTGTGTGAACGCTGTGTGTCCAGTACAAAGCATC 300

QY 311 CCGTGCCCGGACCGGCTCATGATCAGGTGCTAGCTTCTGTGAGGGCAAGATGAGAA 370
Db 301 CCGTGCCGACGACGAGCTGATGTGCTATCTCGGCTAGCTGACGCGGTGAGCAAGCGG 360
QY 371 GGGCGTTTGTGCGGCGGACGCTATCCGCTACGCCATCTGTGGCCAAAGTCTCATCTG 430
Db 361 GGGCGGCTGTGCGCCGACCCCTCATCCGTACCGGAACCTGGCGTCCGCTGTGCTG 420
QY 431 CGGAGCATGACGACCTCGGTATACAGGGCTTTCCACTCTTCAACACCTGTGTATCA 490
Db 421 CGCTCGGTACGACCCCGGTGCTTAAAGGCTTCCACACATGAGACAGTGTGTGACGCA 480
QY 491 GGTTTTATGACCCATGAGGAAACATAGAGAGTTGAGAAAGTTGGGCTCTACACAAACAA 550
Db 481 GGTTCATGTCCCAAGAAAGAGAAAGAAAGTTTGAAGACCTGAAATCCGACTTCAACAG 540
QY 551 TTCTGGGTGCGCTGAGGTGTGTTGCAACTTGTCAATGAAGGCTTATCTTGAAGTCA 610
Db 541 TACTGGGTCCCTGCTGTGTGTACCACTGGCGGCGCCGAGGCGGAGGAGCGGCGA 600
QY 611 ATCCGGACACCGTCTGCTCTCAAGCTGATGATAGGTGTGTACTTTCGTAATCAG 670
Db 601 ATACGTGACGATATCGCTCTCTGTCTACTTGTGAAAGAGCTGAACAACTACCGAGCCA 660
QY 671 TGTGACAGCTGTATGCTTACGACTGATGATATGATATCCATGTGTGTACACAGGTGTG 730
Db 661 TGCAGCATGCTATTCACATAGACTGATGATGATCCCTCTGCTTACACCAAGTGTG 720
QY 731 ACAGTGCAGATATACAGCTTTTCTTGTGATGCTTGTATCGGAGGCACTTTCTGAACCA 790
Db 721 ACCATAGCGGTCTACTCTTCTTGTGCTCTCTGCTGTGCGCCAGTTGTGAGGCA 780
QY 791 AACAAAGACTTACCGAGCCATGATGATGTGTT----- 826
Db 781 GAGGCGAGGGCTGCAAACTCAGAAAGCTTCTGAAGCCAGGCGAGAGCCAGCCCAAGC 840
QY 827 -----GTGCTGTCTTTCACAAATCTGCAATTTCTTATCTACATG 865
Db 841 CTGGAAGACCGGACATGATGATGCTCTCACTGCTGCTGAGTTCTTCTTATGCT 900
QY 866 GCGTGGCTGAAGGTGAGGAGAACAGCTATCAACCTTCGAGGAGAGAGATGATGATTT 925
Db 901 GCGTGGCTGAAGGTGAGGAGAACATATCAACCATTTGTGTGAGATGATGAGACTTT 960
QY 926 GAGACTAATGATCATTTGACAGAAACCTGAGGTGCTCTGTTGCTCGGTGATGGATG 985
Db 961 GAGCAATATGACTCATAGACCGCACTTGAAGGTGTCTGCTATTCGTGAGAGAAATG 1020
QY 986 CACGAAACTTGCTCCCATGAGAACGTGACATGTATGGAAGAGGAGGAGGCTCAGCGG 1045
Db 1021 TACCAAGACTTTCCTCCCGCTGAGAAAGAACAGTATGAGTGAAGACACCGAGCCA 1080
QY 1046 CCTTACAC---AGTGTCTTTCAGAGTCTGCGCGGATTCCTTCAATGAGGCTTCA 1102
Db 1081 CCTTACACTGTGAGGAGGAGGCGGAGTCTGTGGGCTTATTCGTGGGCTTCACTTC 1140
QY 1103 AACATCAGCTTAAAGAAAGA 1122
Db 1141 AACCTGGCATGAGACGAGCA 1160
RESULT 14
ABLI0793
ID ABLI0793 standard; cDNA; 2861 BP.
XX ABLI0793;
XX AC
XX DT 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 26861.
XX DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW

KW pharmaceutical; gene; ss.
 XX Drosophila melanogaster.
 XX WO200171042-A2.
 XX
 XX 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US009231.
 XX
 XX 23-MAR-2000; 2000US-0191637P.
 XX 11-JUL-2000; 2000US-00614150.
 XX
 XX (PEXE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX MPI. 2001-656860/75.
 XX P-PSDB; ABB68690.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signaling and cell-cell
 XX interactions.
 XX
 XX Claim 1; SEQ ID NO 26861; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 XX useful in developmental biology and in elucidating cell signalling and
 XX cell-cell interactions in higher eukaryotes for the development of
 XX insecticides, therapeutics and pharmaceutical drugs. The invention
 XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 XX sequences (AB101840-AB16175) and the encoded proteins (AB57737-
 XX AB872072). The sequence data for this patent did not form part of the
 XX printed specification, but was obtained in electronic format directly
 XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 2861 BP; 761 A; 701 C; 723 G; 676 T; 0 U; 0 Other;
 XX
 XX Query Match 15.5%; Score 296.2; DB 4; Length 2861;
 XX Best Local Similarity 55.4%; Pred. No. 1.9e-64;
 XX Matches 668; Conservative 0; Mismatches 513; Indels 24; Gaps 4;
 XX
 XX 7 AGCCATGACTATCACTTACACAAACAAAGTAGCCA--ATGCCGCGCTCGGTTGTTCTC 63
 XX 411 AACATGACAAATTAAGTACACAGGAGTAGGAGCCACTGTGCGGCTTGGCTGTTTCT 470
 XX
 XX 64 GTCCCTCCTCTGCTGCTGCGGAGAGGAGATCTACAGCTGCTGATGAGAAATCTTGT 123
 XX 471 CAAATGCTGCTCAGATGCGAGAGGAAAGATTACAAACTGCTTTGGCTTGAATCTTGGC 530
 XX
 XX 124 CTTGATATTCCTCTACTATTCATCCGCGAGCTACAGAAATGCTCTGAGATATCA 183
 XX 531 CTTTGTGACATTTATCTATGCGATCAATGATGATGCTTTGGCTCAACCCCGACA 590
 XX
 XX 184 GCAGCTGTTGTTGAGAGCTGCTCTGTAAGCTGAGACGATCAATTCAGCTCATCCCTAT 243
 XX 591 AAAAGAAACCTTTAGAGCCATTTGTTCACTGATGATGATTAACAGAACTCATACCCCT 650
 XX
 XX 244 ATCCCTGCTGCTGCTTCTATGTTTACATTTGCTGATGAGCGCTGCTGAGACCAATGCA 303
 XX 651 GTCCCTGCTGCTGCTTCTATGTTTATGATGATGATGATGATGATGATGATGATGATGAT 710
 XX
 XX 304 GAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
 XX 711 CTCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770
 XX
 XX 364 TGAGAAAGCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
 XX 771 TGAGCGAGAGCGATGATGAGGCGAAACATATGATGATGATGATGATGATGATGATGAT 830
 XX
 XX 424 CATGCTGCGAGCATGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483

Db 831 GATCCTGGCAGATGTTTCGCCGAGGATGAGAGAGCGTTTCCCGGCTTAATATCTGCT 890
 Qy 484 GCTAGCAGGTTTATGACCCATGGGGAACATA-----AGCAGTGCAGAGTTGGG 534
 Db 891 GGAAGCGGCTCTGCTAATGACATGAAAGAACCATCATGAGCCATGAAACAGGCTT 950
 Qy 535 CCTACACACAAACATCTTGGTGGCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 594
 Db 951 TCCAGACCTTTCGAAAGCATGGCTGGCCATGCTTTGGGCTGGCGATTTTAAACGAGGC 1010
 Qy 595 CTATCTTGAAGTGAATTCGGGACACCGCTCTCTCTCAAGCTGATGATGAGTGTG 654
 Db 1011 CAGAAAGAAAGTGCATCTGCTGATGATTTTGTCTGTAAGCAATCATCATGATGCTTAA 1070
 Qy 655 TACTTGGTACTCAGTGTGAGACCTGTATGCTTACAGCTGATGATGATGATGATGATGATGAT 714
 Db 1071 TAACTTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1130
 Qy 715 GTACACACAGGTGTGACAGTGGCAGTATACAGCTTTTCTTTCATGCTGATCGGAG 774
 Db 1131 GTACACCCAGTGTGACCTGGCGGTGTATGCTGATCTTCTTACCTGCTGATGAGTCA 1190
 Qy 775 GCACTTCTGAACCCAA-----CAAGACTACCCAGGCTGATGATGATGATGATGATGAT 825
 Db 1191 ACATGACCCGATGCGAAGGTGTGAGCAATACCATACCTGAAACAAAGTGTGATCTATA 1250
 Qy 826 TGTGCTGTCTTCAATTCCTGCAATTCCTATTTCTACATGAGGCTGAGGAGTGCAGA 885
 Db 1251 CTTTCTCTGTTATTTACACGCTGACAGTTCTTCTTCTACATGAGGTTGCTCAAGTGGCCGA 1310
 Qy 886 ACAGCTCATCAACCCCTTCGGGAGAGAGCATGATGATTTTGAAGTAACTGATCATTTGA 945
 Db 1311 GTGCTGATTAATTCATTTGGCGAAGCATGATGATTTTGAAGTCAACTGATGAGTGA 1370
 Qy 946 CAGAAACCTGCAAGTGTCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1005
 Db 1371 TCGCAATCTTCAGGTGCTCTATCTGATGCTGCAAGATGACATGATGATGATGATGATGAT 1430
 Qy 1006 GAAACGTGACATGATCTGGAAGAGGAGGAGGCGC---TCAGCCGCTTACACAGCTGCTTC 1062
 Db 1431 GTTAAAGATCATGATCTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1490
 Qy 1063 TGCAGGCTTCGCGGATTCCTTCAATGAGGCTTCACCTTCAACATCAAGCTTAAAGAA 1122
 Db 1491 CGAACGATTCGGAAGATCATCAAGGCGCTCAGTCCAGATGAGGAGGAGGAGGAGGAGGAG 1550
 Qy 1123 AGACTTGAAGCTTTGCTCAAAAGAGAGGCTGACAGATTAAGAAAGAGTGGCTATAG 1182
 Db 1551 TCGGCGCATGCTGACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1610
 Qy 1183 CAGCA 1187
 Db 1611 TGGCA 1615
 XX
 XX RESULT 15
 XX AAS72796
 XX ID AAS72796 standard; cDNA; 1526 BP.
 XX
 XX AAS72796;
 XX
 XX 13-FEB-2002 (first entry)
 XX
 XX DNA encoding novel human diagnostic protein #8600.
 XX
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 XX
 XX WO200175067-A2.
 XX
 XX 11-OCT-2001.
 XX PD

XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 PI Drmanac RT, Liu C, Tang YT;
 XX WPI, 2001-639362/73.
 DR P-PSDB; ABG08609.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 PT
 XX
 XX Claim 1; SEQ ID NO 8600; 103bp; English.
 PS
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AA64197-AA694564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 CC
 XX
 XX

Sequence 1526 BP; 404 A; 408 C; 357 G; 357 T; 0 U; 0 Other;

Query Match 13.6%; Score 260.4; DB 5; Length 1526;
 Best Local Similarity 63.7%; Pred. No. 1.8e-55;
 Matches 396; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 492 GTTTTATGACCATGGGAGATGATGACAGTTCAGAGTGGGCTTACCAACACAT 551
 DB 1 GTTTTATGACCAAGATGAAAGAAATTAATCAACCACTCAAGTCTCTCATGGAAT 60
 QY 552 TCTGGGTGCTGGGTGTGTTTGGCACTTGTCAATGAAGGCTTATCTTGAGGTGAA 611
 DB 61 ATTGGGTTCATTATCTGTTTGGAAATCTTGCAATAAGCCCGAATGAAGGTAGAA 120
 QY 612 TCCGGGAACCGTCTGCTCCAGAGCCGTGAATGAGGTGTGTAATTGTGTAATCTAGT 671
 DB 121 TCAGAGACAGTGTGATCTGATCATCTGATGACTGAATGATGATGATACCGCTTTGAT 180
 QY 672 GTGACAGCTGTATGCTACAGACTGTATGATATCCATTGGTGTACACAGAGTGTGA 731
 DB 181 GCAGCTCTTATGCTGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 732 CAGTGGCAGTATACAGCTTTTCTTGTGATGCTTGTGAGGAGGAGTTCGAACCAA 791
 DB 241 CTCTGCTGTCTATCTTCTTCTTGTGCTGATGATGAGCAGGATTTTGTGATCCA 300
 QY 792 ACAAGACTATCCAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 851
 DB 301 CCAAGGCTACGACAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 852 TCTTATCTACATGGGCTGTGTAAGTGGAGAACAGCTCATCAACCTTCGGGGAGG 911

DB 361 TCTTCTTATGACGATGCTTAAGTATGACAGACGACTTATCAACCTTTTGGAGAA 420
 QY 912 ACGATGATGATTTTGAAGCTAATGATGATGATGATGATGATGATGATGATGATGATGAT 971
 DB 421 ATGATGATGATTTTGAAGCTAATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 972 CCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1031
 DB 481 CTGTGAAGAAATGACATGAGCTTACCAAGATGAAAGAGACATTTACTGGAGCAT 540
 QY 1032 CAGGCTCTAGCCGCCCTTACACAGCTCTTGTCCAGGTCTGCGCGGATTCCTTCATGG 1091
 DB 541 CTGCTGCTCGCCCAACATACATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 1092 GCTCCACTTCATCATGACCT 1113
 DB 601 GGTCAACAGTCAAGATGGGCT 622

Search completed: March 25, 2005, 23:52:02
 Job time : 1007.27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2005, 07:39:17 ; Search time 1067.29 Seconds
(without alignments)
10698.156 Million cell updates/sec

Title: US-09-622-964A-28

Perfect score: 1916
Sequence: 1 gtgcacagcatgcatcata.....aaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5552208 seqs, 297965951 residues

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	867	45.3	1758	US-10-492-032-1	Sequence 1, Appl1
2	519.4	27.1	2137	US-10-492-032-3	Sequence 3, Appl1
3	519	27.1	1530	US-10-492-032-30	Sequence 30, Appl1
4	458.2	23.9	2500	US-10-108-260A-1299	Sequence 1299, Ap
5	423	22.1	2028	US-10-492-032-5	Sequence 5, Appl1
6	399.2	20.8	1422	US-10-492-032-7	Sequence 7, Appl1
7	227.6	11.9	1263	US-09-746-783-3	Sequence 3, Appl1
8	194.4	10.1	620	US-09-814-353-19504	Sequence 19504, A
9	187.8	9.8	1350	US-09-768-826-16	Sequence 16, Appl1
10	187.8	9.8	1350	US-10-874-484-16	Sequence 16, Appl1
11	181.2	9.5	1292	US-10-198-846-11070	Sequence 11070, A

C	12	125.2	6.5	578	13	US-10-027-632-134530	Sequence 134530,
C	13	125.2	6.5	578	17	US-10-027-632-134530	Sequence 134530,
C	14	122.2	6.4	539	16	US-10-029-386-7583	Sequence 7583, Ap
C	15	119.8	6.3	235	16	US-10-029-386-21283	Sequence 21283, A
C	16	118	6.2	2914	18	US-10-723-860-7441	Sequence 7441, Ap
C	17	118	6.2	2914	18	US-10-723-860-8303	Sequence 8303, Ap
C	18	114.4	6.0	853	14	US-10-198-846-11346	Sequence 11346, A
C	19	113.8	5.9	1954	18	US-10-723-860-7576	Sequence 7576, Ap
C	20	113.2	5.9	748	13	US-10-027-632-134529	Sequence 134529,
C	21	113.2	5.9	748	17	US-10-027-632-134529	Sequence 134529,
C	22	94.4	4.9	462	14	US-10-198-846-2561	Sequence 2561, Ap
C	23	94.4	4.9	615	14	US-10-198-846-7843	Sequence 7843, Ap
C	24	94.2	4.9	507	10	US-09-764-872-195	Sequence 195, App
C	25	75.8	4.0	3186778	13	US-10-027-632-134961	Sequence 134961,
C	26	75.8	4.0	3186778	17	US-10-027-632-134961	Sequence 134961,
C	27	69	3.6	1717	17	US-10-264-237-259	Sequence 259, App
C	28	66	3.4	239	18	US-10-637-855-110	Sequence 110, App
C	29	66	3.4	65	10	US-09-908-975-29941	Sequence 29941, A
C	30	63.6	3.3	442	10	US-09-814-353-14343	Sequence 14343, A
C	31	63	3.3	751	14	US-10-198-846-1862	Sequence 1862, Ap
C	32	62.2	3.2	260	10	US-09-814-353-1601	Sequence 1601, Ap
C	33	62.2	3.2	260	10	US-09-814-353-7959	Sequence 7959, Ap
C	34	61.6	3.2	298	10	US-09-814-353-806	Sequence 806, App
C	35	61.6	3.2	298	10	US-09-814-353-7178	Sequence 7178, Ap
C	36	60.4	3.2	463	10	US-09-814-353-13563	Sequence 13563, A
C	37	56.2	2.9	3490	9	US-09-925-301-44	Sequence 44, Appl1
C	38	55.6	2.9	440	17	US-10-424-599-5477	Sequence 5477, Ap
C	39	54.6	2.8	644	9	US-09-984-245-84	Sequence 84, Appl1
C	40	54.6	2.8	644	10	US-09-966-262-84	Sequence 84, Appl1
C	41	54.6	2.8	644	10	US-09-983-966-84	Sequence 84, Appl1
C	42	54.6	2.8	644	14	US-10-059-395-84	Sequence 84, Appl1
C	43	54.6	2.8	644	14	US-10-143-090-84	Sequence 84, Appl1
C	44	54.6	2.8	1465	17	US-10-264-237-468	Sequence 468, App
C	45	54	2.8	1872	18	US-10-357-930-24484	Sequence 24484, A

ALIGNMENTS

RESULT 1
Sequence 1, Application US/10492032
Publication No. US20050049212A1
GENERAL INFORMATION:
APPLICANT: Steuermagel, Arnd
APPLICANT: Brommer, Gunter
APPLICANT: Fritsch, Rudiger
APPLICANT: Eulenberg, Karsten
TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
TITLE OF INVENTION: Regulation of Energy Homeostasis
FILE REFERENCE: 2923-606
CURRENT APPLICATION NUMBER: US/10/492,032
CURRENT FILING DATE: 2004-04-08
PRIOR APPLICATION NUMBER: PCT/EP02/11321
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: EP01124059.5
PRIOR FILING DATE: 2001-10-09
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1
LENGTH: 1758
TYPE: DNA
ORGANISM: Homo sapiens
US-10-492-032-1
Query Match 45.3%; Score 867; DB 13; Length 1758;
Best Local Similarity 77.1%; Pred. No. 2.66-23;
Matches 1110; Conservative 0; Mismatches 220; Indels 39; Gaps 3;
QY 11 ATGACTATCACTTACACAAAGTACCAATCCGCGCTTCTGCTCCCTC 70
DB 1 ATGACATCACTTACACAAAGTACCAATCCGCGCTTCTGCTCCCTC 60

QY	71	CTCTGCGTGGGAGGACACATCTAACAACCTCTGATGAGAAATTCCTGTCTTCAAT	130
Db	61	CTCTGTGCTGGGCGGGCAGCATCTCAACAACCTCTAATATGGCAATGTTTAAATCTTCCG	120
QY	131	TTTCTACTATTTCCATCCGTGACATCTACAGATGGTCTCTCGAGTATCGACAGCTG	190
Db	121	CTCTGCTACTAATCAATCCGCTTTATTTAATAGCTGGCCTTACAGAAACACAGCTG	180
QY	191	TTTGTTTGAAAGCTGGCTGTGTACTGTGGACGCTAACATTCACTCATTCCTTAATCTTC	250
Db	181	ATGTTTGAGAAATGTACTGTGATTTGACACAGCTACATCCAGCTCATATCCCATTTCTTC	240
QY	251	GTTCTGGGTTTCTATGTTACATGTTGGTATGCTGAGCGCTGGGGAGCCAGTACAGAACTTG	310
Db	241	GTCGTGGGCTTCTACGTGACGCTGTGTGTGACCCGCTGGTGGAAACATGACGAACCTTG	300
QY	311	CCGTGGCCCGACCGCTCTATGATCCAGTGTCTAAGCTTGTGGAGGGCAAGATGAGAA	370
Db	301	CCGTGGCCCGACCGCTCTATGAGCTCGAGTGTGGGCTTGTGTCGAAGGCAAGGACGAA	360
QY	371	GGCGCTTGTCTGGGGGAGCGCTCATCCGCTACGCAATCTGGGCCAATGCTCATCTCG	430
Db	361	GGCGGCTGTGTGGGGGAGCGCTCATCCGCTACGCAATCTGGGCCAATGCTCATCTCG	420
QY	431	CGCAGCATCAGACCTCGGTCTAACAGCGCTTTCACACTTTCACCACTGGTGTAGCA	490
Db	421	CCAGCGTCTAGACCGCGACGCTTACAAAGCGCTTCCCGAGGCCGACGACTGGTGTAGCA	480
QY	491	GGTTTATATACCCATATGGGGAACATATAGCATTTGACGAATGGGCTTACCAACACACA	550
Db	481	GGCTTATATACCTCCGCAACACAAAGAGTTGGAAAACTAGGCTTACCAACAACTAG	540
QY	551	TTCTGGGTGCTCGGGGTGTGTGTGGCAACTTGTCAATAAGGCTTATTTGAGGTGCA	610
Db	541	TTCTGGGTGCTCGGGGTGTGTGTGGCAACTTGTCAATAAGGCTTATTTGAGGTGCA	600
QY	611	ATCCGGGACACCGTCTGTCTCCAGAGCTGATGAATGAGTGTGTACTTTGGTACTG	670
Db	601	ATCCGGGACACCTATCTGTCTCCAGAGCTGTGAACGAGATGAACACTTGGTGTACTG	660
QY	671	TGTGACAGCTGTATGCTCAGACATGGATTAATATCCATTGGTATACACACAGGTGTG	730
Db	661	TGTGACACCTGTATATGCTTACGACTGTGAATTAATATCCATGTGTATACACAGGTGTG	720
QY	731	ACAGTGGCAGTATACAGCTTTTCTCTTGATGCTTATTCGGAGGACGTTTCTGAACCA	790
Db	721	ACTGTGGCGGTGTACAGCTTCTCTCTGACTTGTCTAGTTGGCGGCACTTCTGAACCA	780
QY	791	AACAGGACCTACCCAGGCCATGATGATCTGTGTGTGCTGTCTTACAAATCTGTCAA	850
Db	781	GCCAAAGCCCTACCTGGCCATGATGCTGTGACCTGTGTGGCCGTTTACAGTTCCTGACG	840
QY	851	TTCTTATTTCTACATGGGCTGTGTGAAGTGTGCAGAAACGCTCATCAACCCCTCGGGAG	910
Db	841	TTCTTCTTCTATGTTGGCTGGGTGAAGTGTGAGAGCACTCATTAACCCCTTTGGAGAG	900
QY	911	GACCATGATGATTTTGAACATACCTGATCATTTGACAGAAACCTGACAGTGTCCCTGTG	970
Db	901	GATGATGATGATTTTGAACACCAATCTGATGTGTGACAGAAATTTGACAGTGTCCCTGTG	960
QY	971	TCCGTGATGATGATGACCAAGAACTTGGCTCCCATGAAACGTGACATTACTGGAAAGAG	1030
Db	961	GCTGTGATGATGATGACCAAGAACTTGGCTGTGATGAGCCGACATTTACTGGAAATAG	1020
QY	1031	GCAGGCTTCAAGCGCCCTTACACAGCTGTTCTGCAAGTGTGCGCGCATTTCTTATG	1090
Db	1021	CCCGAGCCACAGCCCCCTTACACAGCTGTTCCGCGCAGTTCCGTCGAGCTCTCTTATG	1080
QY	1091	GGCTTCAACCTTCAACATCAGGCTTAAAGAAATAAATCTTAAGCTTTGTGCTAAAGAGAG	1150
Db	1081	GGCTTCAACCTTCAACATCAGGCTTAAACAAAGAGATGAGGTTTCAAGCCCAATTCAGAG	1140

QY	1151	GCTGACAGATTAAGAAAGAGAGGCGATATAGCAGACCAATAGGCTGCTTTAGGACTG	1210
Db	1141	G-----ACGAGAGGATGCTCACGCTGGCATTTGGCCGCTTCTTAGGGCTG	1188
QY	1211	CAACCCAAAATTAACAATCTTCCCTTGAAGACTTAAAGACCAACTATTGTGTTCTTAAG	1270
Db	1189	CAGTCCCATGATCAACATCTCTCCAGGGCAAACTCAAGACCAAACTACTGTGGCCAAAG	1248
QY	1271	A-----ACCCCTCTCTCGAAGGCGCATGTATAGATCCCAACGAAAAACCTGAAG--	1322
Db	1249	AGGGATTCCTTCTTCAAGAGGGCTCTGCCAAAAACAACAAGGCAAGCAACGAAACGTT	1308
QY	1323	-----ATGCTGGAATTTTAAGGGCTGTGACTTCTTGAATGTGTT	1363
Db	1309	AGGGGCCAGGAAGACAACAAGGCTGTGAAGCTTGAGCGCCTTCAAGTGTGCC	1368
QY	1364	CCAAGGTTTAAGAGGAAGGCTCCCATTTGTGGCCCAACAGGACCCAGACCAACCTTAC	1422
Db	1369	CCACTGTATCAGAGCCAGGCTACTTAAGTGGCCCAACAGACACCCCTCAGACCCCACTCC	1427

RESULT 2

```

US-10-492-0323
; Sequence 3, Application US/10492032
; Publication No. US20050049212a1
; GENERAL INFORMATION:
; APPLICANT: Steuernagel, Arnd
; APPLICANT: Bronner, Gunter
; APPLICANT: Fritsch, Rudiger
; APPLICANT: Eulenbergh, Karsten
; APPLICANT: Closssek, Thomas
; TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
; TITLE OF INVENTION: Regulation of Energy Homeostasis
; FILE REFERENCE: 2923-606
; CURRENT APPLICATION NUMBER: US/10/492,032
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: PCT/EP02/1121
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: EP01124059.5
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 2137
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-492-032-3

```

Query Match	27.1%	Score 519.4	DB 19	Length 2137
-------------	-------	-------------	-------	-------------

Best Local Similarity 66.9%; Pred. No. 3.2e-135;

Matches 755; Conservative 0; Mismatches 371; Indels 3; Gaps 1;

QY	9	CCATAGCATATACCTACCTACACCAACAAAGAGACCAATGCCCGCTCGGTTCGTTCGTGCC	68
Db	210	CGATGACCCTACCTACCAAGCCGAGGGGAAAGCCCGCTTCGGTGGCTTCTCCAGC	269
QY	69	TCTCTCTGTGCTGGCGAGGACAGACTCTACAAGCTGCTGTATGAGAAATTCCTGTCTTCA	128
Db	270	TCTCTGCTACTGGGCGTGGAGACATCTACAATCTCTGTGGCGAGACTGCTCTGCTTCC	329
QY	129	TATTCCTCTACATTTCCATCCGTGTGACTCTACAAATGCTTCTCTCGAATGATCAGAC	188
Db	330	TTCGGTCTTCACTGGCGCTGAGTGTGTGCTTACCCGCTTTGTGTGACCCGAAAGGCGAAGC	389
QY	189	TGTTGTTTGAAGAGCTGGCTCTGTACTGCGACAGACTACATTTACGCTCATCCCTATATCT	248
Db	390	GCTACTCGAAGAGCTTGTGATTTATTTGTGACACGATATGCGACGCTCATCTCGTCTCGT	449
QY	249	TGCTTCCTGGGTTTCTATATGTTACATTTGTGTGAGCCCTGCTGTGAGCCAGTACGAGAACT	308
Db	450	TGCTCTCTGGCTTTTATATGTAGCGCTGTGTGTGAACCGCTGCTGTGAGCCAGTACTCTATGCA	509
QY	309	TGCCCTGGCCGACCGCCTCATGATCCAGGTGCTTACTCTTGTGAGAGCCAGATGAGG	368

Query	Best Local Similarity	Matches	Score	DB	Length	Indels	Gaps
911 GACGATGATGATTTTGA	63.5%	700	458.2	17	2500	0	0
901 GACGATGATGATCTTTG	63.5%	700	458.2	17	2500	0	0
971 TCCGTGATGGATGACCA	63.5%	700	458.2	17	2500	0	0
961 GCAGTGAGCAGATGATG	63.5%	700	458.2	17	2500	0	0
1031 GCAGGGCCCTCAGCCG	63.5%	700	458.2	17	2500	0	0
1021 GCCGAGGCTGCGCCCA	63.5%	700	458.2	17	2500	0	0
1088 ATGGAGCTCCACCTTC	63.5%	700	458.2	17	2500	0	0
1081 CAGGGCTCCACCTTGA	63.5%	700	458.2	17	2500	0	0
US-10-108-260A-1299	63.5%	700	458.2	17	2500	0	0
Sequence 1299, Application US/10108260A	63.5%	700	458.2	17	2500	0	0
Publication No. US20040005560A1	63.5%	700	458.2	17	2500	0	0
GENERAL INFORMATION:	63.5%	700	458.2	17	2500	0	0
APPLICANT: HELIX RESEARCH INSTITUTE	63.5%	700	458.2	17	2500	0	0
TITLE OF INVENTION: No. US20040005560A1 full length cDNA	63.5%	700	458.2	17	2500	0	0
FILE REFERENCE: H1-A0106	63.5%	700	458.2	17	2500	0	0
CURRENT APPLICATION NUMBER: US/10/108,260A	63.5%	700	458.2	17	2500	0	0
CURRENT FILING DATE: 2002-03-27	63.5%	700	458.2	17	2500	0	0
NUMBER OF SEQ ID NOS: 5458	63.5%	700	458.2	17	2500	0	0
SOFTWARE: Patentin Ver. 2.1	63.5%	700	458.2	17	2500	0	0
SEQ ID NO 1299	63.5%	700	458.2	17	2500	0	0
LENGTH: 2500	63.5%	700	458.2	17	2500	0	0
TYPE: DNA	63.5%	700	458.2	17	2500	0	0
ORGANISM: Homo sapiens	63.5%	700	458.2	17	2500	0	0
US-10-108-260A-1299	63.5%	700	458.2	17	2500	0	0
Query Match	23.9%	700	458.2	17	2500	0	0
Best Local Similarity	63.5%	700	458.2	17	2500	0	0
Matches 700; Conservative 0; Mismatches 403; Indels 0; Gaps 0;	63.5%	700	458.2	17	2500	0	0
11 ATGACTATCACCCTACAA	63.5%	700	458.2	17	2500	0	0
229 ATGACTGCACTTACTCCA	63.5%	700	458.2	17	2500	0	0
71 CTCCTGTCGTCGAGGAG	63.5%	700	458.2	17	2500	0	0
289 CTCCTCAAGTGAAGAGC	63.5%	700	458.2	17	2500	0	0
131 TTCCTCTACTATTCCTC	63.5%	700	458.2	17	2500	0	0
349 GTTCTTTATACAGAAAT	63.5%	700	458.2	17	2500	0	0
191 TTGTTTGAAGAGCTGGC	63.5%	700	458.2	17	2500	0	0
409 TACTTTGAAAAATTAAT	63.5%	700	458.2	17	2500	0	0
251 GTTCTGGATTTCTATGT	63.5%	700	458.2	17	2500	0	0
469 GTGCTGGGTTTATATG	63.5%	700	458.2	17	2500	0	0
311 CCGTGGCCCGACCGCCT	63.5%	700	458.2	17	2500	0	0
529 CCGTGGCCCGACCGCCT	63.5%	700	458.2	17	2500	0	0
371 GGCAGTTTGGCTGGGGCA	63.5%	700	458.2	17	2500	0	0
589 GGGGCGCTGCTTGAAGA	63.5%	700	458.2	17	2500	0	0
431 CGGAGATCAACAAGCTG	63.5%	700	458.2	17	2500	0	0
649 CGCTCGGTGAGCACTGT	63.5%	700	458.2	17	2500	0	0
491 GGTTTATGACCAATGGGA	63.5%	700	458.2	17	2500	0	0

Db	Query Match	Score	DB	Length	2028
Db	709 GGGTTTATGACAAACAGATGAAAGAAATTTATTCACCACTCAAGTCTCTCTCATCTGAAA	62.2%	Pred. No.	4.7e-108	
Qy	551 TTCTGGGTGCTCTGGGTGTGTTGGCCAACTGTCAATGAAAGCTTATCTTGAGGTGCA	62.2%	Pred. No.	4.7e-108	
Db	769 TATTTGGTTTCCATTCACTGCTTTGGAATCTTGCAACCTTAAGCCCGGAATGAAGGTAGA	62.2%	Pred. No.	4.7e-108	
Qy	611 ATCCGGGACACCGGCTCTGCTCCGAGACCTGATGAAATGAGGTGTGTACTTTGCCGATCAG	62.2%	Pred. No.	4.7e-108	
Db	829 ATCAGAGACAGTGTATCTGCATATCTGATGATCTGAAATGAAATGATACCGCTCTTGG	62.2%	Pred. No.	4.7e-108	
Qy	671 TGTGACACAGCTGTATGCTTACGACTGATTAAGTATCCCATTTGGTGTACACAGAGTGTG	62.2%	Pred. No.	4.7e-108	
Db	889 TGGAGCTCTTATATGCGTTATGACTGGGTGGGATTCCGCTGGTTTACACCCAGTTGTCT	62.2%	Pred. No.	4.7e-108	
Qy	731 ACAGTGGCAGTATACAGCTTTTTCCTTGTGATCTGATCGGAGGCGAGTTTCTGAACCCA	62.2%	Pred. No.	4.7e-108	
Db	949 ACTCTTGCTGTCTAATACCTTCTTCTTGGGTGCTATTTGGAAGCCAGTTTGTGATCCC	62.2%	Pred. No.	4.7e-108	
Qy	791 AACAAAGACTACCCAGGCGCATGAGATGATCTGTGTGCTGTCTTCAACATCTGCA	62.2%	Pred. No.	4.7e-108	
Db	1009 ACCAAAGGCTATGCGAGGCGATGACTTGGATCTTTAATCTTCCATCTTCAACCTCTACAA	62.2%	Pred. No.	4.7e-108	
Qy	851 TTCTTATTTCACTAGGCTGCTGCTGAAAGTGCGAACAAGCTCATCAACCCCTTGGGAG	62.2%	Pred. No.	4.7e-108	
Db	1069 TTCTTCTTCTATGACAGATGGCTTAAGTAGACAGACACTTATCAACCTTTTGGAGAA	62.2%	Pred. No.	4.7e-108	
Qy	911 GACGATGATGATTTTGAGCTAAGCTGATCATTTGACAGAAACCTGAGAGTCCCTGTG	62.2%	Pred. No.	4.7e-108	
Db	1129 GATGATGATGATTTTGAACCTAAGCTGCTGATTTGACAGAAATTTGAGAGTCTCTTTT	62.2%	Pred. No.	4.7e-108	
Qy	971 TCCGTGATGGGATGACACCAAGCTTCCCTCCATGGAACGTGACATGTACTGGAACGAG	62.2%	Pred. No.	4.7e-108	
Db	1189 GCTGTGACAGAAATGACATGAGCTTACCAAGATGAAGAGACATTTACTGGAGCAT	62.2%	Pred. No.	4.7e-108	
Qy	1031 GCAGGCGCTCAGGCGCCCTTACACAGCTGTTTCTGCCAGGTCTGCGCGGATTCCTTCAG	62.2%	Pred. No.	4.7e-108	
Db	1249 TCTGCTGCTCGCCACCAATACATGATGGAGCTGTGACTGACTGCAATACCTTCATTTCTG	62.2%	Pred. No.	4.7e-108	
Qy	1091 GGCTCCACCTTCACATCAGCCT	62.2%	Pred. No.	4.7e-108	
Db	1309 GGGTCAACAGTCCAGATGGGCT	62.2%	Pred. No.	4.7e-108	
RESULT 5					
US-10-492-032-5					
; Sequence 5, Application US/10492032					
; Publication No. US20050049212A1					
; GENERAL INFORMATION:					
; APPLICANT: Steuernagel, Arnd					
; APPLICANT: Brommer, Gunter					
; APPLICANT: Fritsch, Rüdiger					
; APPLICANT: Eulenbergh, Karsten					
; APPLICANT: Ciosek, Thomas					
; TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the					
; TITLE OF INVENTION: Regulation of Energy Homeostasis					
; FILE REFERENCE: 2923-606					
; CURRENT APPLICATION NUMBER: US/10/492,032					
; PRIOR FILING DATE: 2004-04-08					
; PRIOR APPLICATION NUMBER: PCT/EP02/11321					
; PRIOR FILING DATE: 2002-10-09					
; PRIOR APPLICATION NUMBER: EP01124059.5					
; PRIOR FILING DATE: 2001-10-09					
; NUMBER OF SEQ ID NOS: 37					
; SOFTWARE: PatentIn version 3.2					
; SEQ ID NO 5					
; LENGTH: 2028					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
US-10-492-032-5					

Matches 697; Conservative 0; Mismatches 400; Indels 24; Gaps 1;

QY 11 ATGACTATACCTACACAAAGTAGCCCAATGCCCGCTGGTGGTCTCGCCCTC 70
Db 1 ATGACTGTCACTTACTCCAGTAAGTAGCAATGCAACTTTTGGATTTTCATAGGTTA 60

QY 71 CTCCTGTGCTGGCAGGAGCAGATCTCAAGCTGTGTATGAGAAATTCCTTGTCTTCATA 130
Db 61 CTCCTCAAGTAGGAGGAGCAGATCTCAAAATCTAGTACAGGAAATTTATTTGTTTGCT 120

QY 131 TTCTCTACTATTCATCCGTGAGCTTACAGAAATGTTCTCTGAGATACAGAGCTG 190
Db 121 GTTCTTTATACAGCAATTAAGTTGGTATACAGATTGTACTTACAGAGGCTCAAAAACGT 180

QY 191 TTGTTTAGAAGCTGGCTGTGACTGAGCAGCTACATTCAGCTCATCCCTATATCCCTC 250
Db 181 TACTTTGAAAATTAATCAATTACTGTACAGATATGCTGAACAAATTCAGTAACCTTT 240

QY 251 GTTCTGGGTTTCTATGTTACATTTGGTGTGAGCCGCTGGTGAAGCAGTACGAGAACTTG 310
Db 241 GTGCTTGGGTTTATGTTACTCTGTGTATGAAACGATGTGTGAACAGTTGTGAAATTTG 300

QY 311 CCGTGCCCGAACCCGCTCATGATCCAGGTGTCTAGCTTCGTGAGGGCAGAGATGAGAA 370
Db 301 CCCTGGCAGACAGGCTAATGTCTCATCTCTAGAGTGTTCACGGAAGCAGAGAGCAC 360

QY 371 GGCCTTTGCTGGCGCGCAGCTCATCCGCTAACCCATCCCTGGGCGAAGTCTCATCTG 430
Db 361 GGGGGCCGCTTTAAGAGACGCTGATGCTACGTCATCTCACCCTCCCTCTCATCTTT 420

QY 431 CGCAGCATCAGACACCTCGGTCTACAGAGGCTTTCCCACTCTTCCACCACTGGTGTGACA 490
Db 421 CGCTCGGTGAGACATCGCTGTGTACAAAGATTTCCCAACATGACACAGCTGGTGAAGCA 480

QY 491 GGGTTTATGACCCATGGGGAACATTAAGCATTTGAGAAATTTGAGGCTTACACACACACA 550
Db 481 GGGTTTATGACAAAGATGAAGAAATTTATCAACCACTCAAGTCTCTCATCTGAAA 540

QY 551 TTCTGGGTCCTGGGTGTGTTGGCCAACTTGTCAAGAAAGGCTTATCTTGAGGTGCA 610
Db 541 TATTTGGGTTCCATTCATCTGTTTGGAAATCTTGCACTAAAGCCGGGAATGAAGGTGA 600

QY 611 ATCCGGGACACCGTCTGCTCCAGAGCCTGATGATGATGATGATGATGATGATGATGATG 670
Db 601 ATCAAGAGCATGTGTGATCTGCAATCATTTGATGATGATGATGATGATGATGATGATG 660

QY 671 TGTGAGCAGCTGTATGCTCAAGCATGATGAATATCCCATTTGTGTATGACACAGGTGTC 730
Db 661 TGACGCTCTTATTCGTTATGACGTGGGTGGGAATCCGCTGGTTTACACCGAGGTGTC 720

QY 731 ACAGTGGCAGTATACAGCTTTTCTGTCATGCTTGTGATGCGGAGGCAAGTTCTGAACCA 790
Db 721 ACTCTTGTCTGTATACCTTCTTCTTGTGCGCTGATGAGACGCAATTTTGTGATCCC 780

QY 791 AACAAAGCTACCAAGGCGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 850
Db 781 ACCAAAGGCTACGAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 840

QY 851 TTCTTATTTCTACATGGGCTGGCTGAG-----GTGGCAGAA 886
Db 841 TTCTTCTTCTTATGACAGATGGCTTAAGGATTAATAATATGTAATCAAGCTAGTGAAG 900

QY 887 CAGCTCATCAACCCCTTGGGGAAGACATGATGATTTTGAAGCTTAACTGATCATTTGAC 946
Db 901 CAGCTTATCAACCCCTTGGGGAAGATGATGATGATTTTGAACCTTAACTGATGATGAC 960

QY 947 AGAAACCGGAGGTGCTGCTGTGTGCGTGTGATGATGATGATGATGATGATGATGATGATG 1006
Db 961 AGAAATTTGAGAGTCTCTTTTAACTGTGAGCAAAATGCAATGAGCTTAACTCAAGATG 1020

QY 1007 GAACGTGACATGTATGGAAGAGGACGCTCAAGCCGCTTACACAGCTGCTTCTGCC 1066
Db 1021 AAGAAAGCATTTTACTGGGAGCATTTCTGTGCTGCGCCACCATTAACATTTGGCAGCTCT 1080

QY 1067 AGGTCTGCCGGCATTCCTTCATGAGCTTCGACCTTCAACAT 1107
Db 1081 GACTACTGCATATCCCTCATTTTCTGGGTCAAGACTCCAGAT 1121

RESULT 6
US-10-492-032-7
; Sequence 7, Application US/10492032
; Publication No. US20050049212A1
; GENERAL INFORMATION:
; APPLICANT: Steuernagel, Arnd
; APPLICANT: Brömer, Günter
; APPLICANT: Fritsch, Rüdiger
; APPLICANT: Eulenber, Karsten
; APPLICANT: Closesek, Thomas
; TITLE OF INVENTION: Bestrophin and Bestrophin Homologous Proteins Involved in the
; FILE REFERENCE: 2923-606
; CURRENT APPLICATION NUMBER: US/10/492,032
; PRIOR FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: PCT/EP02/11321
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: EP01124059.5
; NUMBER OF SEQ. ID NOS: 37
; SOFTWARE: PatentIn version 3.2
; SEQ. ID NO 7
; LENGTH: 1422
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-492-032-7

Query Match 20.8%; Score 399.2; DB 19; Length 1422;
Best Local Similarity 61.1%; Pred. No. 2e-101;
Matches 709; Conservative 0; Mismatches 403; Indels 48; Gaps 2;

QY 11 ATGACTATACCTACACAAAGTAGCCCAATGCCCGCTGGTGGTCTCGCCCTC 70
Db 1 ATGACGGTTTATATCACTCTCAAGTGGCGGAGCCGCTTGGAGTTTCTTGGCCTG 60

QY 71 CTCCTGTGCTGGCAGGAGCAGATCTCAAGCTGTGTATGAGAAATTCCTTGTCTTCATA 130
Db 61 CTCTCGCGTGGAGGGAAGATCTCAAGCTCTCTTCAAGAAATTCCTCTTGGG 120

QY 131 TTCTCTACTATTCATCCGTGAGCTTACAGAAATGTTCTCTGAGATACAGAGCTG 190
Db 121 GCTTGTACGCTGTGCTTACATCCTACCGCTGCTGTGACCCAGAGCAGAGTAC 180

QY 191 TTGTTTAGAAGCTGGCTGTGACTGAGCAGCTACATTCAGCTCATCCCTATATCCCTC 250
Db 181 GTGATGTCAAGGGCCCGTACTGCAACGCTCAGACACTCATTTCTGTGCTTT 240

QY 251 GTTCTGGGTTTCTATGTTACATTTGGTGTGAGCCGCTGGTGAAGCAGTACGAGAACTTG 310
Db 241 GTATGGGTTTCTATGTGATCTCTGTGTGAACGCTGTGTGTGCTCCAGTACAAAGCATC 300

QY 311 CCGTGCCCGAACCCGCTCATGATCCAGGTGTCTAGCTTCGTGAGGGCAGAGATGAGAA 370
Db 301 CCGCTGCAGACAGCAGTGTGATGCTGATCTCGCTAGCGGTGACCGGCTGGACACAGCG 360

QY 371 GGCCTTTGCTGGCGCGCAGCTCATCCGCTAACCCATCCCTGGGCGAAGTCTCATCTG 430
Db 361 GGGCGCTGCTGGCGCGACCTTATCCGCTTACCGGAACCTGGGCTCGGTGCTGCTG 420

QY 431 CGCAGCATCAGACACCTCGGTCTACAGAGGCTTTCCACTTTCACCACTGGTGTGACA 490
Db 421 CGCTCGGTGAGACACCGCGTGTCTTAAGCGCTTCCCAACATGAGACAGTGTGAGAGCA 480

QY 491 GGGTTTATGACCCATGGGGAACATTAAGCATTTGAGAAATTTGAGGCTTACACACACA 550
Db 481 GGGTTTATGTCCTCAGGAAGAGGAAAGATTTGAGAGCCTGAATCCGACTTCAACAG 540

QY	551	TTCTGGGGGACCCCTGGGAGTGGTTGGTTCGCAACTTGTCAATGAAGAGCCCTATCTTGGAGGTGCA	610
Db	541	TACTGGGTCCTCCCTGGCTCTGGTTACCAACTGGGCGCCAGGCCCGAGGGAGCGGGCGA	600
QY	611	ATCCGGGACACCGCTCTGCTGCTCCAGAGCCCTGATGATGAGTGTGTACTTTCGTACTCAG	670
Db	601	ATATGGTGAAGATATGCTCTCTGTGTACTTTTGGAAAGAGCTGAACAACTACCGAGCCAAAG	660
QY	671	TGTGGACAGCTGTATGCTTACGACTGGATTAAGTATCCATTGGTGTATCACACAGTGTGTG	730
Db	661	TGCAGCACTATTCTCACTATGACTGATGATCAGCAATCCCCCTCGTCTACACCAGTGTGTG	720
QY	731	ACAGTGGAGATATCAGCTTTTTCCTTSCATAGCTTGATCGGAGGCAATTCTGAACCCA	790
Db	721	ACCATAGCCGTCTACTCTTCTCTTGGCCCTCTCCCTGTTGGCCGCCAGATTGTGAGCCA	780
QY	791	AACAAGAGCTATCCGAGGCCATGATGATGATCTGGTT-----	826
Db	781	GAGGCAAGGGGGCTGCCAACACTCAGAAAGTTCTGAAGCCAGGCCAGAGCCGCCGCC	840
QY	827	-----GTGCTGTCTTTCACAATCTGCAATTCTTATCTCATG	865
Db	841	CTGGGAGACCCGAGACATGTATGCTGCTCTCACACTGTGCTGAGATTCTTCTCTANCT	900
QY	866	GGCTGTGCTGAAGGTGGCAGAAACAGTATCAACCCCTTCGGGAGAGACGATGATTTT	925
Db	901	GGCTGGCTCAAGGTGCTGAAACAGATCATCAACCCATTGGTGTGAGGATGATGACGACTTT	960
QY	926	GAGACTAATGATCATATGACAGAAACCTGAGAGGTGTCCCTGTGTCCGTGGATGGAGTG	985
Db	961	GAGACAATATCAGCTATATGACCGCAACTTGCAGAGTGTCCCTGTATCCGTGTGAGCAAAAT	1020
QY	986	CACCAAGAACTTGCTTCCCATGGAACGTGACATGTATCTGGAACGAGGACGCGCTTCAGCCG	1045
Db	1021	TACCAAGAACTTCCCCCGCTGAGAAAGGACCAATCTGTGGATGTAGAACCAAGCCGACGCA	1080
QY	1046	CCCTTACAC--AGTGTCTTTCGCCAGAGTCTGCGCGGCAATTCCTTATATGGGCTTCACCTTC	1102
Db	1081	CCCTACACTGTGGCCACGCGCGCGAGTCTGTGGGCCCTCATTTCTGGGCTTCACACTTC	1140
QY	1103	AACATCAGCTTAAGAAAGA	1122
Db	1141	AACCTGCGCATGAGCGACGA	1160
RESULT 7			
US-09-746-783-3			
; Sequence 3, Application US/09746783			
; Publication No. US20030044935A1			
; GENERAL INFORMATION:			
APPLICANT: Jacobs, Kenneth			
McCoy, John M.			
Lavallie, Edward R.			
Racie, Lisa A.			
Treachy, Maurice			
Spaulding, Vikki			
Agostino, Michael J.			
Howes, Steven H.			
Reichel, Kim			
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES			
ENCODING THEM			
NUMBER OF SEQUENCES: 231			
CORRESPONDENCE ADDRESS:			
ADDRESS: Genetics Institute, Inc.			
STREET: 87 Cambridgepark Drive			
CITY: Cambridge			
STATE: MA			
COUNTRY: U.S.A.			
ZIP: 02140			
COMPUTER READABLE FORM:			
MEDIUM TYPE: floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			

```

/      / SOFTWARE: Patentin Release #1.0, Version #1.30
/      / CURRENT APPLICATION DATA:
/      / APPLICATION NUMBER: US/09/746,783
/      / FILING DATE: 21-Dec-2000
/      / CLASSIFICATION: <Unknown>
/      / ATTORNEY/AGENT INFORMATION:
/      /   NAME: Milasincic, Debra J.
/      /   REGISTRATION NUMBER: 46,931
/      / TELECOMMUNICATION INFORMATION:
/      /   TELEPHONE: (617) 227-7400
/      /   TELEFAX: (617) 742-4214
/      / INFORMATION FOR SEQ ID NO: 3:
/      / SEQUENCE CHARACTERISTICS:
/      /   LENGTH: 1263 base pairs
/      /   TYPE: nucleic acid
/      /   STRANDEDNESS: double
/      /   TOPOLOGY: linear
/      / MOLECULE TYPE: cDNA
/      / ;
/      / SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-746-783-3

Query Match          11.9%; Score 227.6; DB 10; Length 1263;
Best Local Similarity 67.7%; Pred.No. 4.8e-53;
Matches 383; Conservative 0; Mismatches 144; Indels 39; Gaps 3;

QY      884 GAACAGCTCATCAACCCCTTGGGGAAGAGCATTATGTTTGAGACTTAATGTGATCTT    943
Db       1 GAGCAGCTCATCAACCCCTTGGAGAGATATGATGATTTTGGACCAACTGATTC    60
QY      944 GACAGAAACCTCGACAGTGTCCCTGTGTCCTGATGGATGACACCAACCTTGCCTCC    1003
Db       61 GACAGGAATTTGCAAGTGTCCCTGTGCTGTGATGATGATGACACAGACTGCTCGG    120
QY      1004 ATGAAACGTGACATGTA CTGGAACGAGCAGCGCTCA GCCGCTTACACAGCTGCTTCT    1063
Db      121 ATGAGGCCGAGCATGTACTGTGAATAAGCCGAGCCACAGCCCCCTTACA CAGCTGCTCC    180
QY      1064 GCCAAGTCTGCCCGCATTTCTTATGAGGCTTCACCTTCAACATTCAGCTTAAAGAAAGA    1123
Db      181 GCCCAGTTCCTTCAGAGCCTCCTTATGAGGCTTCACCTTCAACATTCAGCTTAAAGAG    240
QY      1124 GACTTAGAGCTTGTGTCAAAAGAGAGAGGCTGACACGGATAGAAAGAGTGGCTATGAC    1183
Db      241 GAGATGAGTTCACAGCCCAATCAGGAGG-----ACGAGAGAGATGCTCACGCT    288
QY      1184 AGCACCATTAGGCTCTCTTGAAGACTGCAACCAAAAACTACATCTTCCCTTGAAGAC    1243
Db      289 GGCAATCATTTGGCCGCTTCTAGGCTGCAATGCCATGATCAACATCTCCACAGGCAAAAC    348
QY      1244 TTAAAGACCAAATATTTGTGTTTAAAGAA-----CCCCCTCTCTGAAGGCACTGTAAAG    1297
Db      349 TCAAGAGACCAAACTACTGTGGCCCAAGGAGGAAATCCCTTCTCCAGAGGGCTGCCCAA    408
QY      1298 GATGCCAACCAAGAAAACCAAGAAAT-----GTCTGGAATTT    1336
Db      409 AACCAACAAGCAGCCAACACAGAAAGTTAGGGGCCAGAGAGACACAAAGGCTGGAGATT    468
QY      1337 AAGGATCTGAGACTTCTTGAATGTGTTCCAAAGTTTAAAGAGAGAGGCTCCCATTTGTGGC    1396
Db      469 AAGGCTGTGAGCGGCTTCAAGTCTGCCCACATGTATTCAGAGGCCAGGCTACTACAGTGC    528
QY      1397 CCACAGGCAACCAGCAGCAGCCACTTAC    1422
Db      529 CCACAGAGCGCCCTCAGCCCACTCC    554

RESULT 8
US-09-814-353-19504/c
; Sequence 19504, Application US/09814353
; Publication No., US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
```



```

1  APPLICATION:  Sh1 et al.
2  TITLE OF INVENTION:  18 human secreted proteins
3  FILE REFERENCE:  PFI21P1
4  CURRENT APPLICATION NUMBER:  US/09/768,826
5  CURRENT FILING DATE:  2001-01-25
6  PRIOR APPLICATION NUMBER:  PCT/US00/22350
7  PRIOR FILING DATE:  2000-08-15
8  PRIOR APPLICATION NUMBER:  60/148,759
9  PRIOR FILING DATE:  1999-08-16
10 NUMBER OF SEQ ID NOS:  61
11 SOFTWARE:  Patentin Ver. 2.0
12 SEQ ID NO 16
13
14 LENGTH: 1350
15
16 TYPE:  DNA
17
18 ORGANISM:  Homo sapiens
19
20 FEATURE:
21
22 NAME/KEY:  SITE
23 LOCATION:  (1135)
24
25 OTHER INFORMATION:  n equals a,t,g, or c
26
27 NAME/KEY:  SITE
28 LOCATION:  (1148)
29
30 OTHER INFORMATION:  n equals a,t,g, or c
31
32 NAME/KEY:  SITE
33 LOCATION:  (1166)
34
35 OTHER INFORMATION:  n equals a,t,g, or c
36
37 NAME/KEY:  SITE
38 LOCATION:  (1174)
39
40 OTHER INFORMATION:  n equals a,t,g, or c
41
42 NAME/KEY:  SITE
43 LOCATION:  (1181)

```

```

; NAME/KEY: SITE
; LOCATION: (1209)
; OTHER INFORMATION: n equals a,t,g, or c

```

```

1      LOCATION: (12329)
2      OTHER INFORMATION: n equals a,t,g, or c
3      NAME/KEY: SITE
4      LOCATION: (1266)
5      OTHER INFORMATION: n equals a,t,g, or c
6      NAME/KEY: SITE
7      LOCATION: (1285)
8      OTHER INFORMATION: n equals a,t,g, or c
9      NAME/KEY: SITE
10     LOCATION: (1287)
11     OTHER INFORMATION: n equals a,t,g, or c
12     NAME/KEY: SITE
13     LOCATION: (1290)
14     OTHER INFORMATION: n equals a,t,g, or c
15     NAME/KEY: SITE
16     LOCATION: (1295)
17     OTHER INFORMATION: n equals a,t,g, or c
18     NAME/KEY: SITE
19     LOCATION: (1305)
20     OTHER INFORMATION: n equals a,t,g, or c
21     NAME/KEY: SITE
22     LOCATION: (1324)
23     OTHER INFORMATION: n equals a,t,g, or c
24     NAME/KEY: SITE
25     LOCATION: (1339)
26     OTHER INFORMATION: n equals a,t,g, or c
27     NAME/KEY: SITE
28     LOCATION: (1341)
29     OTHER INFORMATION: n equals a,t,g, or c
30     NAME/KEY: SITE
31     LOCATION: (1343)

```

```

? INFO/VAL : C14
? LOCATION: (1345)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1348)
? OTHER INFORMATION: n equals a,t,g, or c

```


NAME/KEY: SITE
LOCATION: (1348)
OTHER INFORMATION: n equals a,c,t,g, or c
US-10-874-484-16

Query Match 9.8%; Score 187.8; DB 18; Length 1350;
Best Local Similarity 60.1%; Pred. No. 8.3e-42;
Matches 377; Conservative 0; Mismatches 202; Indels 48; Gaps 2;

```
QY 544 CAACACATTCTGGGGTCCCTGGGTGGTGGTCCCACTGTCATGAAAGGCTTACTTGG 603
DB 58 CAACAGACTGGGGTCCCTGGGTGGTGGTCCCACTGTCATGAAAGGCTTACTTGG 117
QY 604 AGTGTGAATCCGGGACACACCGTCTGCTCCAGAGCCTGATGATGATGATGATGATG 663
DB 118 CGGGGGAATAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 177
QY 664 TACTCACTGAGACACTGATGATGATGATGATGATGATGATGATGATGATGATG 723
DB 178 AGCCAGTGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 237
QY 724 GGTGTGACAGTGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATG 783
DB 238 AGTGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 297
QY 784 GAACCCAAACAAAGACTACCCAGGCGATGATGATGATGATGATGATGATGATGATG 826
DB 298 GAGAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 357
QY 827 -----GTGCTGTCTTCAAACTCTGCAATCTCTCAATCTCTTAT 858
DB 358 CCCAGCCTGGGAGACCCGAGCATGATGATGATGATGATGATGATGATGATGATG 417
QY 859 CTAATAGGCTGGCTGGAAGGTGGGAGAAAGCTATCAACCTTCCGGAGAGAGATGA 918
DB 418 CTAATAGGCTGGCTGGAAGGTGGGAGAAAGCTATCAACCTTCCGGAGAGATGA 477
QY 919 TGATTTGAGACTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATG 978
DB 478 CGACTTTGAGCAATATGCTATGATGATGATGATGATGATGATGATGATGATGATG 537
QY 979 TGGGATGACACGAACTTGCCTCCCATGGAACGTGATGATGATGATGATGATGATG 1038
DB 538 CGAATGATGACGAACTTGCCTCCCATGGAACGTGATGATGATGATGATGATGATG 597
QY 1039 TCAGCCGCGCTTACAC--AGTGTCTTTCGCAAGTCTGCGGCAATCTCTTCA 1095
DB 598 GCAGCCACCTTACACTGTGCGCACGCGCGAGTCTGCGGCGCTTCTGCGGCTC 657
QY 1096 CACCTCAACATCAGCCTTAAGAAAGA 1122
DB 658 CACCTCAACCTGCGCATGACGACGA 684
```

RESULT 11
US-10-198-846-11070
Sequence 11070, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1287, 1289, 1290, 1291, 1292
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

Query Match 9.5%; Score 181.2; DB 14; Length 1292;
Best Local Similarity 61.4%; Pred. No. 5.9e-40;
Matches 291; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

```
QY 255 TGGGTTCTATGTTATGATGATGATGATGATGATGATGATGATGATGATGATG 314
DB 318 TAGGTTTATGTTATGATGATGATGATGATGATGATGATGATGATGATGATG 377
QY 315 GGGCCGACCGCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 374
DB 378 GGGCAGACAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 437
QY 375 GTTGTGCGGCGGACGCTATCCGCTACGCAATCTGAGGCAAGTGTCTATCTGCGCA 434
DB 438 GCTGTCTTGAAGGACGCTGATGATGATGATGATGATGATGATGATGATGATG 497
QY 435 GCATCAGACCTGCTGCTACAAAGGCTTCCCACTTCAACCTGCTGCTACAGGAT 494
DB 498 CGTGAGACTGCTGTGTAAGAAATTTCCCAATGACACAGTGTGTAAGCAGGAT 557
QY 495 TTATGACCATGAGGACATTAAGCATGATGATGATGATGATGATGATGATGATG 554
DB 558 TTATGACACAGATGAAGAAATTTATTAACCACTCAAGTCTCTCATCTGAATAT 617
QY 555 GGGTCCCTGGGTGTGTTGGTCCCACTTGTATGAATGAAGCCTTATCTTGAAGTGA 614
DB 618 GGGTCCATTCATCTGTTGTAATCTTGCACCTTAAGCCCGGAATGAAGTGAATCA 677
QY 615 GGGACACCGCTGCTGCTCAGAGCCTGATGATGATGATGATGATGATGATGATG 674
DB 678 GAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 737
QY 675 GACAGCTGATGCTTACGACTGATGAATATCCATGTTGTTAGACACAGGTG 728
DB 738 GCCTTATTCGTTATGATGATGATGATGATGATGATGATGATGATGATGATG 791
```

RESULT 12
US-10-027-632-134530/c
Sequence 134530, Application US/10027632
Publication No. US2002019837A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134530
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134530
```

```

Query Match      6.5%; Score 125.2; DB 13; Length 578;
Best Local Similarity 85.8%; Pred. No. 2.4e-24;
Matches 139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```

QY 488 GCAGGTTTATGACCCATGCGGAGACATAGAGATTGCGGCTTACACACAC 547
    |||
DB 183 GCAGGCTTTATGACTCCGCGAGAACACAGAGATTGAGAACTGAGCCTACACACAC 124
    |||
QY 548 ACATCTGGGGTCCCTGGGGTGGTTTCCCACTGTCAATGAAGGCTTATCTGGAGGT 607
    |||
DB 123 ATGTTCTGGGGTCCCTGGGGTGGTTTCCCACTGTCAATGAAGGCTTATCTGGAGGT 64
    |||
QY 608 CGAATCCGGGACACCGCTCTGCTCCAGAGCCTGATGAATGAG 649
    |||
DB 63 CGAATCCGGGACACCGCTCTGCTCTCAAGAGCCTGCTGAAGTG 22
    |||
```

```

RESULT 13
US-10-027-632-134530/c
; Sequence 134530, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134530
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134530
```

```

Query Match      6.5%; Score 125.2; DB 17; Length 578;
Best Local Similarity 85.8%; Pred. No. 2.4e-24;
Matches 139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
```

```

QY 488 GCAGGTTTATGACCCATGCGGAGACATAGAGATTGCGGCTTACACACAC 547
    |||
DB 183 GCAGGCTTTATGACTCCGCGAGAACACAGAGATTGAGAACTGAGCCTACACACAC 124
    |||
QY 548 ACATCTGGGGTCCCTGGGGTGGTTTCCCACTGTCAATGAAGGCTTATCTGGAGGT 607
    |||
DB 123 ATGTTCTGGGGTCCCTGGGGTGGTTTCCCACTGTCAATGAAGGCTTATCTGGAGGT 64
    |||
QY 608 CGAATCCGGGACACCGCTCTGCTCCAGAGCCTGATGAATGAG 649
    |||
DB 63 CGAATCCGGGACACCGCTCTGCTCTCAAGAGCCTGCTGAAGTG 22
    |||
```

```

RESULT 14
US-10-029-386-7583/c
; Sequence 7583, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7583
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```

; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALUATE 9.00e-56
; OTHER INFORMATION: NT HIT: G115304094, EVALUATE 1.00e-130
; OTHER INFORMATION: SWISSPROT HIT: O76090, EVALUATE 2.00e-25
US-10-029-386-7583
```

```

Query Match      6.4%; Score 122.2; DB 16; Length 539;
Best Local Similarity 68.4%; Pred. No. 1.6e-23;
Matches 169; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
```

```

QY 247 CTTCGTTCTGGGTTTCTATGTTACATTGGTGTGAGCCGCTGTGAGCCAGTACAGAA 306
    |||
DB 372 CTCCTCCGACGGCTTTATATGACGCTGTGTGGAACCGCTGTGAGCCAGTACCTATG 313
    |||
QY 307 CTTCGCGTGGCCCGACCGCCTCATGATCCAGGTGTCTAGCTTGTGAGGAGCAAGATGA 366
    |||
DB 312 CATGCCCTGCCCGACCGCTCATGTGCGTGTGGGGGACCGCTGTGAGCCAGTACCTATG 253
    |||
QY 367 GGAAGCGGTTTCTGTGGGGGACAGCTCATCCGCTACGCGCATCTGAGGCGCAAGTGTCTAT 426
    |||
DB 252 CCGGGGCGGCTCTACCGGGGACACTCATGCGTGTGGGGGACCGCTGTGAGCCAGTACCTATG 193
    |||
QY 427 CTTCGCGAGCATAGACACTCGGTCTTACAGAGCCTTCCACTCTTCACTCACTGTGTGCT 486
    |||
DB 192 CTTGCGCTCCGTGAGACCGCGGTGTTCAGAGCCTTCCCACTAGACCACTGTGTGGA 133
    |||
QY 487 AGCAGGT 493
    |||
DB 132 GGCTGGT 126
    |||
```

```

RESULT 15
US-10-029-386-21283/c
; Sequence 21283, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21283
; LENGTH: 235
```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALUE 4.00e-56
; OTHER INFORMATION: NT HIT: g18923136, EVALUE 1.00e-129
; OTHER INFORMATION: SWISSPROT HIT: O76090, EVALUE 1.00e-24
US-10-029-386-21283

Query Match 6.3%; Score 119.8; DB 16; Length 235;
Best Local Similarity 69.4%; Pred. No. 4.8e-23;
Matches 163; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 258 GTTCTATGTTACATTGGTGTGAGCCGCTGCTGAGCCAGTACGAGAACTTGCCGTGGC 317
Db 235 GCTTTTATGTAGCGCTGTGTGAAACCGCTGTGTGAGCCAGTACCTATGATGCCGTGC 176
QY 318 CCGACCGCTCATGATCCAGGTGCTAGCTTCTGTGAGGCGAAGATGAGAGCCGTT 377
Db 175 CCGACGCGCTCATGTGCGTGTGCGGACCGTGCAGACGCGACCGCGCGCC 116
QY 378 TGCTGCGGCGCAGCGCTATCCGCTACGCGCATCTGGGCCAAGTGCTCACTCTGCCAGCA 437
Db 115 TCTACCGCGCACACTCATGTGCGCTACGAGGGCTCTCGGCCGTGCTCATCTGCGCTCCG 56
QY 438 TCAGCACCCTGGGCTTACAAAGCGCTTTCCCACTTTTCAACCACTGGTGTCTAGCAGG 492
Db 55 TCAGCACCAGCGGTGTTCAAAGCGCTTCCCAACATAGAACAGTGTGTGAGGCTGG 1

Search completed: March 26, 2005, 22:07:50
Job time : 1075.29 sec8

This Page Blank (usp10)

D	b		 	380	GAGAACCTGCGGTGGCCCAACGCCCTTCATBAGCCTTGTTGGGACTTTCGTCAAAGGCAG	439
O	y		 	362	GATTAGAAAGCCGTTTGCTGCGGCGCACGCTCATCCGCTAACCCATCTCGGGCCAATG	421
D	b		 	440	GACGAACAAGGCGGGCTGCTGGCGCGCAGCTCATCCGCTTAAGCCAACTGGCCAACTGT	499
O	y		 	422	CTCATCTCGCGGACGATCAGACACTGGGTCTAACAAGCGTTTTCCACTTTCACACTG	481
D	b		 	500	CTCATCTCGCGGACGCTCAGCACCGCAGCTTCAACGCGTTTTCCCAGCGCCACGACTG	559
O	y		 	482	GTCGTAGCAGGTTTTATGACCCTCATGGGGAACTATAAGCAGTTTGAGAAGTTGGGCTAACCA	541
D	b		 	560	GTGCAAGCAGCGCTTTATGACTCCGGAGAACTACAAGCAGTTGGAGAAACTGAGCTTACCA	619
O	y		 	542	CACAACACATTCGGGTGCCCCGAGTAGGTGGTTTTGCCAATTGTCAA TGAAGGCTTACTT	601
D	b		 	620	CACAACATGTTCTGGGTGCCCCGAGTAGGTGGTTTTGCCAATTGTCAA TGAAGGCTTACTT	679
O	y		 	602	GGAGGTCTGAATCCGGGACACCGTCTGCTCAGAGCTTGATGATAGATGGTTACTTTG	661
D	b		 	680	GGAGGTCTGAATCCGGGACACCGTCTGCTCAGAGCTTGATGATAGATGGTTACTTTG	739
O	y		 	662	CGTACTCAGTGTGAGCAGCTGTATGCTTGAGCTGGATAGTATGCCATTGGGTATACA	721
D	b		 	740	CGTACTCAGTGTGAGCAGCTGTATGCTTGAGCTGGATAGTATGCCATTGGGTATACA	799
O	y		 	722	CAGGTGTGACAGATGGCAGTATACAGCTTTTTCTTGACATGCTTGAATCGGAGGCACTT	781
D	b		 	800	CAGGTGTGACTGTGGCGGTGTACAGCTTCTTCCTGACTGTCTAGTTGGGGGGCACTT	859
O	y		 	782	CTGAACCCCAAACAAGACTACCAAGGCACTGAGATGATCTGTTGTGCTGTCTTACA	841
D	b		 	860	CTGAACCCCAAACAAGGCTTACCCTGGCCATGAGCTGGAACCTGTGTGCCCCGTTCACG	919
O	y		 	842	ATCTGCAATTTCTTATTACATGGGCTGGCTGAGAGGTGGGAA CAGCTCATCAACCCC	901
D	b		 	920	TTCCTGCACTTCTTCTTATGTGGCTGCTGAGAGSTGGCAGAGCCTCATCAACCCC	979
O	y		 	902	TTCCGGGAGAGCAGATGATATTTTGAGACTAACTGGATCATTGACAGAAACTGCAAGTG	961
D	b		 	980	TTTGAGAGGAGATGATGATATTTTGAGACCACTGGATTTGTGACACAGAAATTTGCAAGTG	103
O	y		 	962	TCCCTGTTGTCCGTGATGGGATGCACTCAACCAATTCGCTCCCATGGACGTGACATGAC	102
D	b		 	1040	TCCCTGTTGTCCGTGATGGGATGCACTCAACCAATTCGCTCCCATGGACGTGACATGAC	109
O	y		 	1022	TGGAACGAGGCGAGCGCTCAGCGCGGCTTACAGAGCTCTTCGCAAGCTCGCGCGCAT	108
D	b		 	1100	TGGAATTAAGCCCGAGCCACAGCCCCCTTACACAGCTGTCTTCGCGCCAGTTCCGTGAGCC	115
O	y		 	1082	TCCCTTCATGGGCTCCACCTTCAACATCAGCCTTAAAGAAAGAAAGACTTAGAGCTTGGTCA	114
D	b		 	1160	TCCCTTATAGGCTCCACCTTCAACATCAGCCTTAAAGAAAGAAAGATGAGTTCCAGCCC	121
O	y		 	1142	AAGAGAGGCTGACACGGATTAAGAAAGAGTGCCTATGACGACCATAGGCTGCTT	120
D	b		 	1220	AATCAGGAGGAGCAGAGAGGATCTCACGCTG-----GCATCATTTGGCGCTTC	126
O	y		 	1202	TTAGGACCTGCAACCCAAAAATACATCTTCCCTTGAAAGACTTTAAAGCCAACTATTTG	126
D	b		 	1268	CTAGGCTGCACTCCCATGATCACATCTTCCCAAGGCAAACTCAAGGACCAAACTACGCT	132
O	y		 	1262	TGTTTCTAAGA-----ACCCCTCTCGAAGGCGCAGTGTAAAGATGGCCAACAGAAAAC	131
D	b		 	1328	TGGCCCAAGAGGAAATCTCTTCTTCCACAGAGGCTCTGCCMAAAACA CAAGGCAAGCABA	138
O	y		 	1316	CAGAAAG-----ATGCTGAGAAATTTAAGAGCTGTGGACTTCTTG	135
D	b		 	1388	CAGAAAGCTTAAAGGAGCAGAAAGAACAAAGAGGCTTAAAGCTTAAAGCTGTGAGCCCTTC	144
O	y		 	1355	AAATGTGTTCCAAGTTTAAAGAGAGAGCTTCCATTTGTGCCCCACAGGACCTCACGACCT	141

Db	1448	AAGTGTGCCCCACTGTATCAGAGGCCAGGCTACTACAGTGGCCCAACAGACGCCCTCAGC	1507
Qy	1415	CACCCCTAC	1422
Db	1508	CCCACTCC	1515

```

RESULT 2
US-09-949-016-16643
: Sequence 16643, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CL001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 16643
: LENGTH: 18075
: TYPE: DNA
: ORGANISM: Human
: US-09-949-016-16643

```

Query Match	9.8%	Score 188.4	DB 4	Length 18075
Best Local Similarity	81.1%	Pred. No. 8.6e-41		
Matches 219	Conservative	0	Mismatches 51	Indels 0
QY	233	CTCATCCCTATATCTTGTGTTCTGAGTTTCTATGTATCATGTGTGTGAGCCGTGTGTG	232	
Db	7315	CTGCCCCCGCCCTCTCTGACCAGAGCTTTACGTGAGCGGTGTCGTGACCCGCTGTGG	7374	
QY	293	AGCCAGTACGAAACTTGTGCGTGTGCGCCGACCCGCTCATGATCCAGGTGTCTAGCTTGTG	352	
Db	7375	AACCAATTAAGAAACCTCCGTGTGCGCCGACCGCTCATGTAGCCGTGTGTGTGCTTC	7434	
QY	353	GAGGGCAAGATGATGAAAGCCGTTTGTGTGCGGCGACGCTCATCCGCTACGCGATCTTG	412	
Db	7435	GAAAGCGAAGAGAGAGCAAGGCCGCTGTGTGCGGCGACGCTCATCCGCTACGCGCAACTG	7494	
QY	413	GGCGAAGTGTCTATCTTGTGCGGAGCATCGACACTCGTGTCTACAAGCGCTTTCACCTTT	472	
Db	7495	GGCAACGTGTCTATCTCTCGCAGCGGTGTGACACCCGACGTACAAAGCGCTTCCCAAGCC	7554	
QY	473	CACCACTGTGTCTAGCAGGTTTATAGCC	502	
Db	7555	CAGCACTGTGTGCAAGCAGGTGGCGAC	7584	

RESULT 3

US-09-949-016-173051

Sequence 173051, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: ParseSeq for Windows Version 4.0
;; SEQ ID NO: 173051
;; LENGTH: 601
;; TYPE: DNA
;; ORGANISM: Human
US-09-949-016-173051

Query Match 6.5%; Score 125.2; DB 4; Length 601;
Best Local Similarity 85.8%; Pred. No. 3.3e-24;
Matches 139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 488 GCAGGTTTATGACCCATGAGGAGGAAACATAGAGTTGCGAGAGTTGGGCTTACACACAC 547
Db 321 GCAGGCTTTATGACTCCCGGACAGAACAGAGCTTTGAGAAACAGGCTTACACACAC 380
Qy 548 ACATTCTGGGTCCTCGGCTGGGTGGTTTCCCACTTGTCAATGAAAGCCCTATCTTGAGGT 607
Db 381 ATGTTCTGGGTGCTCGGCTGGGTGGTTTCCCACTGTCAATGAAAGCGTGGCTTGAGGT 440
Qy 608 CGAATCCGGGACACCGTCTCTGCTCCAGAGCCTGATGAATGAG 649
Db 441 CGAATCCGGGACCCATCTCTGCTCCAGAGCCTGCTGAACGTG 482

RESULT 4
US-09-949-016-173052
; Sequence 173052, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: ParseSeq for Windows Version 4.0
; SEQ ID NO: 173052
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-173052

Query Match 6.5%; Score 125.2; DB 4; Length 601;
Best Local Similarity 85.8%; Pred. No. 3.3e-24;
Matches 139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 488 GCAGGTTTATGACCCATGAGGAGGAAACATAGAGTTGCGAGAGTTGGGCTTACACACAC 547
Db 30 GCAGGCTTTATGACTCCCGGACAGAACAGAGCTTTGAGAAACAGGCTTACACACAC 89
Qy 548 ACATTCTGGGTCCTCGGCTGGGTGGTTTCCCACTTGTCAATGAAAGCCCTATCTTGAGGT 607
Db 90 ATGTTCTGGGTGCTCGGCTGGGTGGTTTCCCACTGTCAATGAAAGCGTGGCTTGAGGT 149
Qy 608 CGAATCCGGGACACCGTCTCTGCTCCAGAGCCTGATGAATGAG 649
Db 150 CGAATCCGGGACCCATCTCTGCTCCAGAGCCTGCTGAACGTG 191

RESULT 5
US-09-270-767-735
; Sequence 735, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

;; APPLICANT: Homburger et al.
;; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;; FILE REFERENCE: File Reference: 7326-094
;; CURRENT APPLICATION NUMBER: US/09/270,767
;; CURRENT FILING DATE: 1999-03-17
;; NUMBER OF SEQ ID NOS: 62517
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 735
;; LENGTH: 419
;; TYPE: DNA
;; ORGANISM: Drosophila melanogaster
US-09-270-767-735

Query Match 5.2%; Score 100; DB 4; Length 419;
Best Local Similarity 57.9%; Pred. No. 2.4e-17;
Matches 197; Conservative 0; Mismatches 140; Indels 3; Gaps 1;

Qy 851 TTCTTAATTCATAGGCTGGCTGAGAGTGAGCAAGCAAGCTCATCAACCCCTTGGGGAG 910
Db 2 TTCTTCTTCAATAGGCTGGCTCAAGGTGAGCGAGTCGATGAATTCATTGGCGAA 61
Qy 911 GACGATGATGTTTGGAGACTAATCTGATCATTTGACAGAACTTCAGAGTTCCTGTG 970
Db 62 GACGATGATGATTTTGGAGTCAACTGATGATGATGATGATGATGATGATGATGATG 121
Qy 971 TCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1030
Db 122 ATGCTGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 181
Qy 1031 GCAGCGCC--TCAGCGCCCTTACAGAGCTCTTCTGCGAGTCTGCGGCAATTCCTTC 1087
Db 182 GTGTTCCCAAGAGAGTGGCTCCCTTACAGATAGCTGCGGAACGATTCGGGAGATCATCA 241
Qy 1088 ATGAGCTTCACTTCAATCATAGCTTAAGAAAGAGCTTGAAGCTTTGTCAAAAG 1147
Db 242 GAGCGCTCACTGCAAGATGAGAGTCCCAAGATGAGGAGGAGGAGGAGGAGGAGGAG 301
Qy 1148 GAGCGTGAACGATTAAGAAAGAGTGGCTATGACAGCA 1187
Db 302 TCCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 341

RESULT 6
US-09-270-767-16017
; Sequence 16017, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 16017
; LENGTH: 419
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16017

Query Match 5.2%; Score 100; DB 4; Length 419;
Best Local Similarity 57.9%; Pred. No. 2.4e-17;
Matches 197; Conservative 0; Mismatches 140; Indels 3; Gaps 1;

Qy 851 TTCTTAATTCATAGGCTGGCTGAGAGTGAGCAAGCAAGCTCATCAACCCCTTGGGGAG 910
Db 2 TTCTTCTTCAATAGGCTGGCTCAAGGTGAGCGAGTCGATGAATTCATTGGCGAA 61
Qy 911 GACGATGATGATTTTGGAGACTAATCTGATCATTTGACAGAACTTCAGAGTTCCTGTG 970
Db 62 GACGATGATGATTTTGGAGTCAACTGATGATGATGATGATGATGATGATGATGATG 121
Qy 971 TCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1030

Db 122 ATGCTCAGCAGATGACACATGACATCCGAGCTGTAAAGATCACTACTGGAGCAG 181
| 1031 GCAGCGCC---TCAGCCGCCCTTACACACAGCTCTTCTGCGAGGTCTCGCCCGCATTCCTTC 1087
Qy 182 GTGTTCCCAACGAGCTGCCCCCTTACACATATGCTCCGAACATTCGCGAATTCACCA 241
| 1088 ATGGGCTCCACCTTCAACATCAGCTTAAAGAAAGACTTGAAGCTTGGTCAAAAGAG 1147
Db 242 GAGCGCTCAGCTGACAGATGAGGTGCCCCAAGATGGCGCATGCCATGACATATGCG 301
Qy 1148 GAGGCTACACGATTAAGAAAGAGTGGCTATTAAGCAGCA 1187
| 302 TCCGTGCACTGATGAATGCGCATGATGCCAGTGGCA 341

RESULT 7
US-09-949-016-173045
; Sequence 173045, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173045
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-173045

Query Match 3.8%; Score 72.2; DB 4; Length 601;
Best Local Similarity 82.2%; Pred. No. 1.4e-09;
Matches 83; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
Qy 2 TGGCAAGCAGTATGATCACTTACCAAGCAAGTAGCCGCTGCTGCTTC 61
| 501 TGCTGGCAGACATCACTTACCAAGCAAGTAGCCGCTTGGCTTC 560
Db 62 TCGTCCCTCTCTGCTGCTGCGAGCGAGCATCTTCAAGCT 102
| 561 TCCGCGCTGCTGCTGCTGCGAGCGAGCATCTTCAAGCT 601

RESULT 8
US-09-949-016-16644/c
; Sequence 16644, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16644
; LENGTH: 7609
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16644

Query Match 3.6%; Score 68.6; DB 4; Length 7609;
Best Local Similarity 58.1%; Pred. No. 5.8e-08;
Matches 175; Conservative 0; Mismatches 99; Indels 27; Gaps 2;

Qy 1149 AGGCTGACAGGATTAAGAAAGAGTGTATAGACACATTAAGCTGTCTTAGAC 128
| 7353 AGCCCAATCAGAGAGAGAGAGATCTACCGCTGACATATTTGGCCCTTCTTAGGCC 7294
Db 1209 TGCAACCCAAAATTAACATCTTCTTGAAGACTTAAGACCAATTAATGTTCTA 1268
| 7293 TGCACTCCATGATACATCTCTCCAGGCGAAACTCAAGAGCCAACTGTGGGCCA 7224
Qy 1269 AGA-----ACCCCTCTCTGAAAGGCGATGTAAGATGCCAACAGAAAAACGAAAG 1322
| 7233 AGAGGAATCCCTTCTCAGAGGGCTGCCCAAAAACCAAGGCGCAAGACAGAACG 7174
Db 1323 AT-----GCTGAAATTTAAGGATCTGGAATTTGGAATGTG 1361
| 7173 TTAGGGGCCAGAAACACACAAAGGCTTGAAAGCTTAAAGCTGTGAGCGCTTCAAGTCTG 7114
Qy 1362 TTCCAAGTTTAAAGAGAGAGGCTCCCATTTGTGGCCACAGAGCAGCCAGCCCTA 1421
| 7113 CCCCACTGTATCAGAGGCCAGGCTACTACAGTGGCCCAAGAGGCCCTCAGCCCACTC 7054
Qy 1422 C 1422
| 7053 C 7053
Db

RESULT 9
US-09-949-016-173053
; Sequence 173053, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173053
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-173053

Query Match 3.2%; Score 61.4; DB 4; Length 601;
Best Local Similarity 91.5%; Pred. No. 1.3e-06;
Matches 65; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 662 CGTACTAGTGTGACAGCTGTATGCTTACGATGATTAAGTATCCATTTGTATACA 721
| 1 CGTACTAGTGTGACAGCTGTATGCTTACGATGATTAAGTATCCATTTGTATACA 60
Db 722 CAGGTGTGAC 732
| 61 CAGGTGAGAC 71

FILE REFERENCE: 07334-076001
CURRENT APPLICATION NUMBER: US/09/099,041A
CURRENT FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 09/019,942
PRIORITY FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1931
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (214)...(1833)
US-09-099-041A-1

Query Match 2.7%; Score 51.8; DB 3; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.0011;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 1814 AAACACCTGATTTTACGATCTTCCCAACTAAGAGTTTAAATGAATATTCTT 1873
Db 1825 AAAAGCATGTAGAGCTGTTTTCACAGAGAAATGTGTTTCATTAAGATATTATTA 1884
Qy 1874 TTAGGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
Db 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927

RESULT 14
US-09-245-281-1
Sequence 1, Application US/09245281
Patent No. 6369196
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
FILE REFERENCE: 07334/118001
CURRENT APPLICATION NUMBER: US/09/245,281
CURRENT FILING DATE: 1999-02-05
EARLIER APPLICATION NUMBER: US 09/207,359
EARLIER FILING DATE: 1998-12-08
EARLIER APPLICATION NUMBER: US 09/099,041
EARLIER FILING DATE: 1998-06-17
EARLIER APPLICATION NUMBER: US 09/019,942
EARLIER FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 44
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1931
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (214)...(1833)
US-09-245-281-1

Query Match 2.7%; Score 51.8; DB 3; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.0011;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 1814 AAACACCTGATTTTACGATCTTCCCAACTAAGAGTTTAAATGAATATTCTT 1873
Db 1825 AAAAGCATGTAGAGCTGTTTTCACAGAGAAATGTGTTTCATTAAGATATTATTA 1884
Qy 1874 TTAGGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
Db 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927

RESULT 15
US-09-470-271-2
Sequence 2, Application US/09470271

Patent No. 6410689
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
TITLE OF INVENTION: DOMAIN POLYPEPTIDES
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/470,271
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,942
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Melk1eJohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/068001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1931 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-470-271-2

Query Match 2.7%; Score 51.8; DB 3; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.0011;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
Qy 1814 AAACACCTGATTTTACGATCTTCCCAACTAAGAGTTTAAATGAATATTCTT 1873
Db 1825 AAAAGCATGTAGAGCTGTTTTCACAGAGAAATGTGTTTCATTAAGATATTATTA 1884
Qy 1874 TTAGGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
Db 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927

Search completed: March 26, 2005, 13:49:44
Job time : 308.441 secs